

STIC-Biotech/ChemLib

98810

Fr m: STIC-ILL
Sent: Tuesday, July 15, 2003 10:12 AM
To: STIC-Biotech/ChemLib
Subject: FW: 09/910582- Search of 5 Peptide SEQ's)+ Interf.DB)

-----Original Message-----

From: Audet, Maury
Sent: Tuesday, July 15, 2003 10:02 AM
T : STIC-ILL
Subject: 09/910582- Search of 5 Peptide SEQ's)+ Interf.DB)

On the above case, could I have you search the following 5 peptide SEQ's (including Interference DB's RAPM, RAPN): 2, 3, 4, 9, 10.

Thanks Maury

Maury Audet
Patent Examiner
Art Unit 1654
703-305-5039
11D04

Edward Hart
Technical Info. Specialist
STIC/Biotech
MAIL ROOM? Tel: 305-9203

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 7/16/03
Searcher Prep/Review: 7/16/03
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: 5
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: DA
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:04:13 ; Search time 23.9024 Seconds
(without alignments)
39.023 Million cell updates/sec

Title: US-09-910-582B-2

Perfect score: 44

Sequence: 1 GGGVFWQ(7)

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002:*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	7	22	Peptide which sele
2	44	100.0	7	22	Heart homing pepti
3	38	86.4	215	22	Novel human diagno
4	38	86.4	328	23	Human cancer cell
5	38	86.4	431	23	Herbicide-activ
6	37	84.1	210	21	Zeas may protein f
7	37	84.1	219	22	Novel human diagno
8	37	84.1	311	21	Zeas may protein f
9	37	84.1	352	21	Zeas may protein f
10	37	84.1	470	20	H. pylori outer me

11	37	84.1	480	19	AAW98219	H. pylori GHPO 170
12	36	81.8	280	21	AAW56841	Human prostase can
13	35	79.5	56	22	AAW89732	Human immune/haema
14	35	79.5	131	23	ABW49927	Listeria monocytog
15	35	79.5	149	22	ABG13868	Novel human diagno
16	35	79.5	210	21	AAW20020	Arabidopsis thalia
17	35	79.5	231	21	AAW20019	Arabidopsis thalia
18	35	79.5	236	21	AAW20018	Arabidopsis thalia
19	35	79.5	236	21	AAW20018	Arabidopsis thalia
20	35	79.5	236	21	AAW20018	Arabidopsis thalia
21	35	79.5	236	21	AAW20018	Arabidopsis thalia
22	35	79.5	236	21	AAW20018	Arabidopsis thalia
23	35	79.5	236	21	AAW20018	Arabidopsis thalia
24	35	79.5	236	21	AAW20018	Arabidopsis thalia
25	35	79.5	236	21	AAW20018	Arabidopsis thalia
26	35	79.5	236	21	AAW20018	Arabidopsis thalia
27	35	79.5	236	21	AAW20018	Arabidopsis thalia
28	35	79.5	236	21	AAW20018	Arabidopsis thalia
29	35	79.5	236	21	AAW20018	Arabidopsis thalia
30	35	79.5	236	21	AAW20018	Arabidopsis thalia
31	35	79.5	236	21	AAW20018	Arabidopsis thalia
32	35	79.5	236	21	AAW20018	Arabidopsis thalia
33	35	79.5	236	21	AAW20018	Arabidopsis thalia
34	35	79.5	236	21	AAW20018	Arabidopsis thalia
35	35	79.5	236	21	AAW20018	Arabidopsis thalia
36	35	79.5	236	21	AAW20018	Arabidopsis thalia
37	35	79.5	236	21	AAW20018	Arabidopsis thalia
38	35	79.5	236	21	AAW20018	Arabidopsis thalia
39	35	79.5	236	21	AAW20018	Arabidopsis thalia
40	34	77.3	81	22	AAW01437	Human protein sequ
41	34	77.3	85	22	AAW11374	Novel human diagno
42	34	77.3	115	21	AAW52862	Arabidopsis thalia
43	34	77.3	140	21	AAW52861	Arabidopsis thalia
44	34	77.3	163	21	AAW52860	Arabidopsis thalia
45	34	77.3	238	21	AAW52860	Pinus radiata cell

ALIGNMENTS

RESULT 1

AAW30895
ID AAW30895 standard; peptide; 7 AA.

XX AAW30895;

XX 02-APR-2001 (first entry)

XX Peptide which selectively binds to normal cardiac endothelium.

XX Cardiac endothelium; angiogenic factor; vascular endothelium;
XX peripheral vascular disease; cardiovascular disease; angiogenesis;
XX cardiac neovascularisation.

OS Unidentified.

XX WO200075329-A1.

PN WO200075329-A1.

XX 14-DEC-2000.

XX 31-MAY-2000; 2000WO-US14988.

XX 07-JUN-1999; 99US-0327045.

XX (EDWA-) EDWARDS LIFESCIENCES CORP.

PA (BAXT) BAXTER AG.

XX Levine AJ, Mitterer A, Falkner F, Scheiflinger F, Dorner F;

XX WPI; 2001-091212/10.

XX New chimeric molecules having an angiogenic factor linked to a

PT targeting molecule that binds to a vascular endothelium, useful for

PT

PT increasing cardiac neovascularisation, or treating peripheral vascular
PT and cardiovascular diseases

PS Disclosure; Page 27; 67pp; English.

XX AAB30895-99 represent targeting molecules, which are used to produce the
XX chimeric molecules of the invention. AAB30895-98 selectively bind to
CC normal cardiac endothelium. The specification describes a chimeric
CC molecule comprising an angiogenic factor linked to a targeting molecule
CC that specifically binds to a vascular endothelium. The chimeric
CC molecules are useful for treatment of peripheral vascular or
CC cardiovascular diseases. Specifically, they are useful for inducing or
CC inhibiting angiogenesis, for increasing cardiac neovascularisation in
CC ischemic tissue in the peripheral vascular system.

XX Sequence 7 AA;

Query Match 100.0%; Score 44; DB 22; Length 7;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVFWQ 7

DB 1 GGGVFWQ 7

RESULT 2

AAB50797

ID AAB50797 standard; peptide; 7 AA.

AC AAB50797;

XX 21-MAR-2001 (first entry)

DE Heart homing peptide SEQ ID NO: 2.

XX Heart homing peptide; cardiovascular disease; ischaemic disease;
KW gene therapy.

XX Synthetic.

OS WO200075174-A1.

PN 14-DEC-2000.

PD 31-MAY-2000; 2000WO-US15088.

PF 07-JUN-1999; 99US-0326718.

PR (BURN-) BURNHAM INST.

XX Ruoslahti E, Mackenna DA;

XX WPI; 2001-071059/08.

XX Novel heart homing peptide that selectively homes to normal ischaemic
PT and cardiac tissue useful for targeting ischaemic tissues for treating
PT ischaemic and cardiovascular diseases such as atherosclerosis and
PT restenosis

PS Claim 2; Page 55; 70pp; English.

XX The present invention provides a number of heart homing peptides which
CC selectively home to cardiac tissue. These can be used in the treatment of
CC cardiovascular and ischaemic diseases, such as atherosclerosis,
CC thrombosis, restenosis, vasculitis, atherosclerotic aneurysms, myocardial
CC hypertrophy, congenital heart diseases, ischaemic heart disease and
CC anginas, acquired valvular/endocardial diseases, primary myocardial
CC diseases, cardiac tumours and arrhythmias.

XX Sequence 7 AA;

Query Match 100.0%; Score 44; DB 22; Length 7;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVFWQ 7

DB 1 GGGVFWQ 7

RESULT 3

ABG19611

ID ABG19611 standard; Protein; 215 AA.

XX AC ABG19611;

XX 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #19602.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS83798.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 20; SEQ ID No 49970; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 215 AA;

Query Match 86.4%; Score 38; DB 22; Length 215;

Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGVFWQ 7
Db 153 GGGVMHQ 159

RESULT 4
ABB09234
ID ABB09234 standard; Protein; 328 AA.
XX ABB09234;
DT 09-JUL-2002 (first entry)
XX Human cancer cell growth promoting protein PP3999 SEQ ID NO:2.
DE Human; cancer cell growth promoting protein; PP3999; PP4534; cytostatic;
XX cancer.
KW Homo sapiens.
OS
XX
XX CN1329065-A.
XX
XX 02-JAN-2002.
XX 20-JUN-2000; 2000CN-0116620.
XX 20-JUN-2000; 2000CN-0116620.
XX (SHAN-) SHANGHAI INST TUMOR.
XX
XX Gu J, Yang S;
XX WPI: 2002-330565/37.
DR N-PSDB; ABL51724.
XX
XX Novel human protein with function of promoting growth of cancer cell
XX useful for treating cancer -
XX
XX Claim 1; Page 13 (Disclosure); 19pp; Chinese.
XX
XX The present sequence represents a human cancer cell growth promoting
CC protein designated PP3999. Also described is a human cancer cell growth
CC promoting protein designated PP4534. PP3999 and PP4534 have cytostatic
CC activities and may be used in the treatment of cancer.
XX
XX
SQ Sequence 328 AA;

Query Match 86.4%; Score 38; DB 23; Length 328;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGVFWQ 7
Db 259 GGAVFWQ 265

RESULT 5
ABB90809
ID ABB90809 standard; Protein; 431 AA.
XX ABB90809;
XX
XX 31-MAY-2002 (first entry)
DT
XX Herbicidally active polypeptide SEQ ID NO 20.
DE
XX Herbicidal; plant; agriculture; herbicide.
XX
XX Arabidopsis thaliana.
OS
XX WO200210210-A2.
PN
XX

PD 07-FEB-2002.
XX
XX 28-AUG-2001; 2001WO-EP09892.
XX
XX 28-AUG-2001; 2001WO-EP09892.
XX
XX (FARB) BAYER AG.
XX
XX Tietjen K, Weidler M;
PI
XX WPI: 2002-269010/31.
DR
XX
XX Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
XX Claim 5; SEQ ID NO 20; 261pp + Sequence Listing; English.
PS
XX
XX The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX
XX SQ Sequence 431 AA;

Query Match 86.4%; Score 38; DB 23; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGVFWQ 7
Db 377 GGVFWQ 382

RESULT 6
AAG33249
ID AAG33249 standard; Protein; 210 AA.
XX
XX AAG33249;
AC
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX Zea mays protein fragment SEQ ID NO: 40257.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
XX Zea mays subsp. mays.
OS
XX
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 99US-0130077.

us-09-910-582b-2.rag

Wed Jul 16 13:41:34 2003

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 84.1%; Score 37; DB 21; Length 210;
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGVFWQ 7
 II:III
 DB 183 GGIFWQ 188

RESULT 7
 ABC01936
 ID ABG01936 standard; Protein; 219 AA.

XX AC ABG01936;
 XX DT 13-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #1927.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS66123.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX PS Claim 20; SEQ ID No 32295; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 219 AA;

Query Match 84.1%; Score 37; DB 22; Length 219;
 Best Local Similarity 71.4%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGVFWQ 7
 III:II

DB 113 GGGAFWE 119

RESULT 8
 AAG33248
 ID AAG33248 standard; Protein; 311 AA.

XX AC AAG33248;

XX DT 18-OCT-2000 (first entry)

XX DE Zea mays protein fragment SEQ ID NO: 40256.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence; corn.

XX OS Zea mays subsp. mays.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

XX PR 19-APR-1999; 99US-0130077.

XX PR 21-APR-1999; 99US-0130449.

XX PR 23-APR-1999; 99US-0130510.

XX PR 28-APR-1999; 99US-0130891.

XX PR 30-APR-1999; 99US-0131449.

XX PR 30-APR-1999; 99US-0132048.

XX PR 30-APR-1999; 99US-0132407.

XX PR 04-MAY-1999; 99US-0132484.

XX PR 05-MAY-1999; 99US-0132485.

XX PR 06-MAY-1999; 99US-0132486.

Wed Jul 16 13:41:34 2003

us-09-910-582b-2.rag

PR	06-MAY-1999;	99US-0132487.	PR	22-JUL-1999;	99US-0145192.
PR	07-MAY-1999;	99US-0132863.	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	99US-0134256.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	99US-0134218.	PR	26-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134219.	PR	27-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134220.	PR	27-JUL-1999;	99US-0145918.
PR	14-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	18-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145951.
PR	19-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146386.
PR	20-MAY-1999;	99US-0135133.	PR	02-AUG-1999;	99US-0146388.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146389.
PR	24-MAY-1999;	99US-0135629.	PR	03-AUG-1999;	99US-0147038.
PR	25-MAY-1999;	99US-0136021.	PR	04-AUG-1999;	99US-0147204.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147302.
PR	28-MAY-1999;	99US-0137282.	PR	05-AUG-1999;	99US-0147192.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147260.
PR	03-JUN-1999;	99US-0137528.	PR	06-AUG-1999;	99US-0147303.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147416.
PR	07-JUN-1999;	99US-0137724.	PR	09-AUG-1999;	99US-0147493.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138540.	PR	10-AUG-1999;	99US-0148171.
PR	10-JUN-1999;	99US-0138847.	PR	11-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139452.	PR	13-AUG-1999;	99US-0148565.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148684.
PR	17-JUN-1999;	99US-0139492.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139454.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139455.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139456.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139459.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139461.	PR	25-AUG-1999;	99US-0150366.
PR	18-JUN-1999;	99US-0139462.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139463.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151066.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151080.
PR	21-JUN-1999;	99US-0139817.	PR	30-AUG-1999;	99US-0151303.
PR	22-JUN-1999;	99US-0139889.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140353.	PR	01-SEP-1999;	99US-0151930.
PR	23-JUN-1999;	99US-0140354.	PR	07-SEP-1999;	99US-0152363.
PR	24-JUN-1999;	99US-0140695.	PR	10-SEP-1999;	99US-0153070.
PR	28-JUN-1999;	99US-0140823.	PR	13-SEP-1999;	99US-0153758.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0154018.
PR	30-JUN-1999;	99US-0141287.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0141642.	PR	20-SEP-1999;	99US-0154779.
PR	01-JUL-1999;	99US-0142154.	PR	22-SEP-1999;	99US-0155139.
PR	02-JUL-1999;	99US-0142055.	PR	23-SEP-1999;	99US-0155486.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155659.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0156458.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156596.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0157117.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157753.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157865.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158232.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159638.
PR	20-JUL-1999;	99US-0144884.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0144814.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160767.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145089.			

PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 84.18; Score 37; DB 21; Length 311;
Best Local Similarity 83.36; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 GGVFWQ 7
||:||||
Db 284 GGIFWQ 289

RESULT 9
AAG33247
ID AAG33247 standard; Protein: 352 AA.

XX AC AAG33247;

XX DT 18-OCT-2000 (first entry)

XX DE Zea mays protein fragment SEQ ID NO: 40255.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX RW hybridisation assay; genetic mapping; gene expression control; promoter;
XX OS termination sequence; corn.

XX OS Zea mays subsp. mays.

XX PN EPI033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.

```
PR 27-JUL-1999; 99US-0145919;
PR 28-JUL-1999; 99US-0145951;
PR 02-AUG-1999; 99US-0146386;
PR 02-AUG-1999; 99US-0146388;
PR 02-AUG-1999; 99US-0146389;
PR 03-AUG-1999; 99US-0147038;
PR 04-AUG-1999; 99US-0147204;
PR 04-AUG-1999; 99US-0147302;
PR 05-AUG-1999; 99US-0147392;
PR 05-AUG-1999; 99US-0147260;
PR 06-AUG-1999; 99US-0147303;
PR 06-AUG-1999; 99US-0147416;
PR 09-AUG-1999; 99US-0147493;
PR 09-AUG-1999; 99US-0147935;
PR 10-AUG-1999; 99US-0148171;
PR 10-AUG-1999; 99US-0148319;
PR 11-AUG-1999; 99US-0148341;
PR 12-AUG-1999; 99US-0148565;
PR 13-AUG-1999; 99US-0148684;
PR 16-AUG-1999; 99US-0149368;
PR 17-AUG-1999; 99US-0149175;
PR 18-AUG-1999; 99US-0149426;
PR 20-AUG-1999; 99US-0149722;
PR 20-AUG-1999; 99US-0149723;
PR 20-AUG-1999; 99US-0149929;
PR 23-AUG-1999; 99US-0149902;
PR 23-AUG-1999; 99US-0149930;
PR 25-AUG-1999; 99US-0150566;
PR 26-AUG-1999; 99US-0150884;
PR 27-AUG-1999; 99US-0151065;
PR 27-AUG-1999; 99US-0151066;
PR 27-AUG-1999; 99US-0151080;
PR 30-AUG-1999; 99US-0151303;
PR 31-AUG-1999; 99US-0151438;
PR 01-SEP-1999; 99US-0151930;
PR 07-SEP-1999; 99US-0152363;
PR 10-SEP-1999; 99US-0153070;
PR 13-SEP-1999; 99US-0153758;
PR 15-SEP-1999; 99US-0154018;
PR 16-SEP-1999; 99US-0154039;
PR 20-SEP-1999; 99US-0154779;
PR 22-SEP-1999; 99US-0155139;
PR 23-SEP-1999; 99US-0155486;
PR 24-SEP-1999; 99US-0155659;
PR 28-SEP-1999; 99US-0156458;
PR 29-SEP-1999; 99US-0156596;
PR 04-OCT-1999; 99US-0157117;
PR 05-OCT-1999; 99US-0157753;
PR 06-OCT-1999; 99US-0157865;
PR 07-OCT-1999; 99US-0158029;
PR 08-OCT-1999; 99US-0158232;
PR 12-OCT-1999; 99US-0158369;
PR 13-OCT-1999; 99US-0159293;
PR 13-OCT-1999; 99US-0159294;
PR 13-OCT-1999; 99US-0159295;
PR 14-OCT-1999; 99US-0159329;
PR 14-OCT-1999; 99US-0159330;
PR 14-OCT-1999; 99US-0159331;
PR 14-OCT-1999; 99US-0159637;
PR 14-OCT-1999; 99US-0159638;
PR 18-OCT-1999; 99US-0159584;
PR 21-OCT-1999; 99US-0160741;
PR 21-OCT-1999; 99US-0160767;
PR 21-OCT-1999; 99US-0160770;
PR 21-OCT-1999; 99US-0160814;
PR 21-OCT-1999; 99US-0160815;
PR 22-OCT-1999; 99US-0160980;
PR 22-OCT-1999; 99US-0160981;
PR 22-OCT-1999; 99US-0160989;
PR 25-OCT-1999; 99US-0161404;
PR 25-OCT-1999; 99US-0161405;
PR 25-OCT-1999; 99US-0161406;

PR 26-OCT-1999; 99US-0161359;
PR 26-OCT-1999; 99US-0161360;
PR 26-OCT-1999; 99US-0161361;
PR 28-OCT-1999; 99US-0161920;
PR 28-OCT-1999; 99US-0161992;
PR 28-OCT-1999; 99US-0161993;
PR 29-OCT-1999; 99US-0162142;

Query Match      84.1%; Score 37; DB 21; Length 352;
Best Local Similarity 83.3%; Pred No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 GGVFWQ 7
DB      325 GGIFWQ 330
      ||:||||

RESULT 10
AAV17215
ID AAV17215 standard; Protein; 470 AA.
XX
AC AAV17215;
XX
DT 03-AUG-1999 (first entry)
XX
DE H. pylori outer membrane polypeptide.
XX
KW Outer membrane polypeptide; OMP; vaccine; H. pylori infection; humoral;
KW cellular immune response.
XX
OS Helicobacter pylori.
XX
PN WO9921959-A2.
XX
PD 06-MAY-1999.
XX
PF 28-OCT-1998; 98WO-US22883.
XX
PR 17-DEC-1997; 97US-0993001.
PR 28-OCT-1997; 97US-0959131.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Alm RA, Ellis RW, Guld BC, Noonan BM, Smith D;
XX
DR WPI; 1999-326698/27.
XX
DR N-PSDB; AAX75834.
XX
PT Cellular vaccine against Helicobacter pylori
PS Claim 7; Page 308-309; 352pp; English.
XX
CC The invention relates to a vaccine for preventing or treating infections
CC by Helicobacter pylori. The vaccine contains at least one isolated
CC H. pylori polypeptide, or its fragments, in a carrier, where the
CC carrier is a Salmonella, Vibrio cholerae or Shigella vector containing a
CC nucleic acid encoding the H. pylori polypeptide. The vaccines induce a
CC humoral and cellular immune responses. The vaccines are used to treat or
CC prevent infections by H. pylori. Sequences AAX75779 to AAX75837 represent
CC nucleic acid sequences encoding H. pylori outer membrane polypeptides
CC (OMPs) AAY17160 to AAY17218.
XX
SQ Sequence 470 AA;

Query Match      84.1%; Score 37; DB 20; Length 470;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 GGVFWQ 7
DB     120 GGIFWQ 126
      ||:||||
```

RESULT 11
AAW98219 standard; Protein; 480 AA.
XX AC AAW98219;
XX DT 31-MAR-1999 (first entry)
XX DE H. pylori GHPO 1706 protein.
XX KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
XX KW peptic ulcer disease.
XX OS Helicobacter pylori.
XX PN WO9843478-A1.
XX PD 08-OCT-1998.
XX PF 01-APR-1998; 98WO-US06371.
XX PR 29-JUL-1997; 97US-0902615.
XX PR 01-APR-1997; 97US-0833457.
XX PR 24-JUN-1997; 97US-0881227.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (INNR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
XX PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;
XX DR WPI; 1998-542293/46.
XX DR N-PSDB; AAX13938.
XX PT New isolated Helicobacter polynucleotides - used to develop products
XX PT for the diagnosis, prevention and treatment of Helicobacter
XX PT infections and gastrointestinal diseases
XX PS Claim 8; Page 141-144; 2054pp; English.
XX CC This sequence represents a Helicobacter pylori GHPO protein of the
XX CC invention. The polypeptides can be used for preventing or treating
XX CC Helicobacter infections, and gastroduodenal diseases associated with
XX CC these infections, including acute, chronic, and atrophic gastritis, and
XX CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
XX CC used for the production of antibodies. The products can also be used for
XX CC detection and diagnosis.
XX SQ Sequence 480 AA;
Query Match 84.1%; Score 37; DB 19; Length 480;
Best Local Similarity 71.4%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GGGVFWQ 7
Db 130 GAGIFWQ 136
RESULT 12
AAB56841 standard; Protein; 280 AA.
XX AC AAB56841;
XX DT 13-MAR-2001 (first entry)
XX DE Human prostate cancer antigen protein sequence SEQ ID NO:1419.
XX KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
XX KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
XX KW vulnary; gastrolntestinal; nephrotropic; antinfctive; gynaecological;
XX KW antibacterial; gene therapy; immune; reproductive; renal;
XX KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;

KW wound; infectious disease.
XX OS Homo sapiens.
XX PN WO200055174-A1.
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WO-US05988.
XX PR 12-MAR-1999; 99US-0124270.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (ROSE/) ROSEN C A.
XX PI Rosen CA, Ruben SM;
XX DR WPI; 2000-587513/55.
XX DR N-PSDB; AAF16044.
XX PT Prostate cancer associated gene sequences, referred to as prostate
XX PT cancer antigens, useful for treatment, prevention, and diagnosis of
XX PT disorders such as prostate cancer -
XX PS Claim 11; Page 1850-1852; 2338pp; English.
XX CC AAF15566 to AAF16505 encode the human prostate cancer associated
XX CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
XX CC The prostate cancer antigens can have neuroprotective, cytostatic,
XX CC cardioactive, immunomodulatory, muscular, vulnary, gastrolntestinal,
XX CC nephrotropic, antinfctive, gynaecological and antibacterial activities,
XX CC and can be used in gene therapy. The prostate cancer antigen
XX CC polynucleotides may be used for detection of prostate cancer, chromosome
XX CC identification, as chromosome markers, and for numerous other diagnostic
XX CC or research purposes. The prostate cancer antigens may be used to treat
XX CC disorders such as neural, immune, muscular, reproductive,
XX CC gastrolntestinal, pulmonary, cardiovascular, renal, and proliferative
XX CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
XX CC AAB57303 represent sequences used in the exemplification of the present
XX CC invention.
XX SQ Sequence 280 AA;
Query Match 81.8%; Score 36; DB 21; Length 280;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGGVFW 6
Db 101 GGGVFW 106
RESULT 13
AAM89732 standard; Protein; 56 AA.
XX ID AAM89732
XX AC AAM89732;
XX DT 07-NOV-2001 (first entry)
XX DE Human immune/haematopoietic antigen SEQ ID NO:17325.
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW cytostatic; gene therapy; vaccine; metastasis.
XX OS Homo sapiens.
XX PN WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01354.
XX

us-09-910-582b-2.rag

Wed Jul 16 13:41:34 2003

PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236329.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PR (HUMA-) HUMAN GENOME SCI INC.
PR Rosen CA, Barash SC, Ruben SM;
PR WPI; 2001-483426/52.
PR N-PSDB; AAK62513.
PR Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PR useful for preventing, diagnosing and/or treating cancers and

PT metastasis -

PS Claim 11; SEQ ID NO 17325; 3071pp + Sequence Listing; English.

XX

CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)

CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic

CC activity, and can be used in gene therapy and vaccine production. (I)

CC proteins and polynucleotides may be used in the prevention, diagnosis and

CC treatment of diseases associated with inappropriate (I) expression. For

CC example, they may be used to treat disorders associated with decreased

CC expression by rectifying mutations or deletions in a patient's genome

CC that affect the activity of (I) by expressing inactive proteins or to

CC supplement the patient's own production of (I). Additionally, (I)

CC polynucleotides may be used to produce the secreted (I), by inserting

CC the nucleic acids into a host cell and culturing the cell to express the

CC protein. (I) proteins and polynucleotides may be used to prevent,

CC diagnose and treat immune/haematopoietic-related diseases, especially

CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703

CC to AAK87694 represent human immune/haematopoietic antigen genomic

CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169

CC represent sequences used in the exemplification of the present invention.

XX

SQ Sequence 56 AA;

Query Match 79.5%; Score 35; DB 22; Length 56;

Best Local Similarity 83.3%; Pred. No. 67;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGVFWQ 7

DB 16 GGVFEW 21

RESULT 14

ABBA4927

ID ABB49927 standard; Protein; 131 AA.

XX

AC ABB49927;

XX

XX

DT 05-FEB-2002 (first entry)

XX

DE Listeria monocytogenes protein #2631.

XX

XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;

KW vitamin B12; bacterial infection; disease.

KW

XX Listeria monocytogenes.

OS

XX WO200177335-A2.

PN

XX

PD 18-OCT-2001.

XX

XX

PF 11-APR-2001; 2001WO-FR01118.

XX

XX

PR 11-APR-2000; 2000FR-0004629.

XX

XX (INSP) INST PASTEUR.

PA

XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;

PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;

PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;

PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;

PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;

PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;

PI Meduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;

PI Rose M, Voss H;

XX

XX WPI; 2002-010914/01.

DR

XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment

PT and prevention of Listeria and related bacterial infections, and

PT related polypeptides -

XX

PS Claim 6; SEQ ID NO 2632; 192pp; French.

XX

CC The present invention relates to the genome sequence of Listeria

CC monocytogenes Ecd-e (see ABA03041). The genome sequence and fragments of

CC it are useful for selecting probes and primers for detecting genes in L.

CC monocytogenes and related organisms, and for studying genetic

CC polymorphisms and other genomes. The present sequence is a protein

CC encoded by the genome sequence of the present invention. Proteins

CC expressed from the genome sequence are useful for raising specific

CC antibodies, identification of L. monocytogenes and related organisms, and

CC for biosynthesis and biodegradation, especially biosynthesis of vitamin

CC B12. The genome sequence and proteins encoded by it are also useful for

CC selecting compounds that regulate gene expression and cell replication

CC and modulate L. monocytogenes-related diseases. In addition, the genome

CC sequence and proteins encoded by it are useful in pharmaceutical and

CC vaccines compositions for the treatment or prevention of infections by L.

CC monocytogenes and related organisms.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 131 AA;

Query Match 79.5%; Score 35; DB 23; Length 131;

Best Local Similarity 83.3%; Pred. No. 1.5e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGVFW 6

DB 67 GGTFW 72

RESULT 15

ABG13868

ID ABG13868 standard; Protein; 149 AA.

XX

AC ABG13868;

XX

XX

DT 18-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #13859.

XX

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

KW

XX Homo sapiens.

OS

XX WO200175067-A2.

PN

XX

PD 11-OCT-2001.

XX

XX

PF 30-MAR-2001; 2001WO-US08631.

XX

XX

PR 31-MAR-2000; 2000US-0540217.

PR

XX 23-AUG-2000; 2000US-0549167.

XX

XX (HYSE-) HYSEQ INC.

PA

XX Drmanac RT, Liu C, Tang YT;

PI

XX WPI; 2001-639362/73.

DR

XX N-PSDB; AAS78055.

DR

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

PT

XX Claim 20; SEQ ID NO 44227; 103pp; English.

PS

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 149 AA;

Query Match 79.5%; Score 35; DB 22; Length 149;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGVFWQ 7
Db 113 GGVFWE 118
|||||

Search completed: July 16, 2003, 13:12:36
Job time : 24.9024 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:10:13 ; Search time 8.19512 Seconds
(without alignments)
25.132 Million cell updates/sec

Title: us-09-910-582b-2

Perfect score: 44

Sequence: 1 GGGVFWQ 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/1aa/PTCUS_COMB.pep.*

6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	7	4	US-09-326-718-2
2	33	75.0	263	4	US-09-134-001C-4512
3	33	75.0	305	3	US-09-178-610-6
4	33	75.0	317	2	US-08-619-362A-8
5	33	75.0	317	2	US-08-790-572-3
6	33	75.0	317	2	US-09-213-398-3
7	33	75.0	318	2	US-08-619-362A-9
8	33	75.0	340	2	US-08-790-572-1
9	33	75.0	340	2	US-09-213-398-1
10	33	75.0	342	4	US-09-149-476-695
11	33	75.0	359	4	US-09-154-750A-90
12	33	75.0	382	2	US-08-872-302-2
13	33	75.0	606	4	US-09-041-236-4
14	32	72.7	81	4	US-09-134-001C-3196
15	32	72.7	291	4	US-09-362-831-4
16	32	72.7	389	1	US-07-939-501A-1
17	32	72.7	389	4	US-08-448-398-7
18	32	72.7	423	1	US-07-939-501A-10
19	32	72.7	423	1	US-07-939-501A-12
20	32	72.7	424	1	US-08-045-269C-2
21	32	72.7	424	3	US-08-371-680-2
22	32	72.7	424	5	PCT-US94-01198-2
23	32	72.7	963	1	US-08-537-002A-3
24	32	72.7	963	3	US-08-863-010-3
25	32	72.7	963	4	US-09-024-429-3
26	32	72.7	1241	4	US-09-040-774-2
27	31	70.5	19	2	US-08-392-816-8

28	31	70.5	24	4	US-09-227-357-661	Sequence 661, Appl
29	31	70.5	187	2	US-08-713-825-1	Sequence 1, Appli
30	31	70.5	187	3	US-09-199-842-1	Sequence 1, Appli
31	31	70.5	364	4	US-09-338-671-2	Sequence 2, Appli
32	31	70.5	383	4	US-09-059-769-12	Sequence 12, Appl
33	31	70.5	383	4	US-09-161-994A-16	Sequence 16, Appl
34	31	70.5	498	1	US-08-357-598-9	Sequence 9, Appli
35	31	70.5	498	2	US-09-003-289-9	Sequence 9, Appli
36	31	70.5	498	5	PCT-US95-16435-9	Sequence 9, Appli
37	31	70.5	573	3	US-09-008-481A-7	Sequence 7, Appli
38	31	70.5	573	4	US-09-195-666A-50	Sequence 50, Appl
39	31	70.5	573	4	US-09-309-592-7	Sequence 7, Appli
40	31	70.5	573	4	US-09-635-705-50	Sequence 50, Appl
41	31	70.5	573	4	US-09-634-858A-50	Sequence 50, Appl
42	30	68.2	116	2	US-08-428-197-10	Sequence 10, Appl
43	30	68.2	116	5	PCT-US93-10555-10	Sequence 10, Appl
44	30	68.2	118	1	US-08-414-926A-3	Sequence 3, Appli
45	30	68.2	118	2	US-08-926-922-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-326-718-2
; Sequence 2, Application US/09326718
; Patent No. 6303573
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; TITLE OF INVENTION: Heart Homing Peptides and Methods of
; FILE REFERENCE: P-LJ 3512
; CURRENT APPLICATION NUMBER: US/09/326,718
; CURRENT FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-09-326-718-2

Query Match	100.0%	Score 44;	DB 4;	Length 7;
Best Local Similarity	100.0%	Pred. No. 2e+05;		
Matches	7;	Conservative	0;	Mismatches
			0;	Indels
			0;	Gaps
			0;	
Oy	1	GGGVFWQ 7		
Db	1	GGGVFWQ 7		

RESULT 2
US-09-134-001C-4512
; Sequence 4512, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4512
; LENGTH: 263
; TYPE: PRT

4.

QY 2 GGVFW 6
| | | | |
Db 97 GGVFW 101

RESULT 6
US-09-213-398-3
; Sequence 3, Application US/09213398
; Patent No. 5955429
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN APOPTOSIS-ASSOCIATED PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/213,398
; FILING DATE: 12/15/98
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/790,572
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0204 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1256606
US-09-213-398-3

Query Match 75.0%; Score 33; DB 2; Length 317;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 2 GGVFW 6
| | | | |
Db 97 GGVFW 101

RESULT 7
US-08-619-362A-9
; Sequence 9, Application US/08619362A
; Patent No. 5843659
; GENERAL INFORMATION:
; APPLICANT: SOPHIE M. LEHAR; and
; APPLICANT: BRAYDON C. GUILD
; TITLE OF INVENTION: NOVEL APOPTOSIS GENE EI24,
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr

STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: District of Columbia
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/619,362A
FILING DATE: 21 MARCH 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WIXON, Henry N.
REGISTRATION NUMBER: 32,073
REFERENCE/DOCKET NUMBER: 104322.170
TELEPHONE: (202) 942-8400
TELEFAX: (202) 942-8484
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 318 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-619-362A-9

Query Match 75.0%; Score 33; DB 2; Length 318;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 2 GGVFW 6
| | | | |
Db 98 GGVFW 102

RESULT 8
US-08-790-572-1
; Sequence 1, Application US/08790572
; Patent No. 5858715
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN APOPTOSIS-ASSOCIATED PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/790,572
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0204 US
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166

Wed Jul 16 13:41:34 2003

us-09-910-582b-2.ra1

TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: 2445356
US-08-790-572-1

Query Match 75.0%; Score 33; DB 2; Length 340;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGVFW 6
Db 79 GGVFW 83

RESULT 9
US-09-213-398-1
Sequence 1, Application US/09213398
Patent No. 5955429
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL HUMAN APOPTOSIS-ASSOCIATED PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/213,398
FILING DATE: 12/15/98
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/790,572
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0204 US
TELEPHONE: 415-855-0355
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: 2445356
US-09-213-398-1

Query Match 75.0%; Score 33; DB 2; Length 340;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGVFW 6

Db 79 GGVFW 83
RESULT 10
US-09-149-476-695
Sequence 095, Application US/09149476
Patent No. 6420526
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149,476
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601

; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 75.0% Score 33; DB 4; Length 342;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGVFW 6
 |||||
Db 81 GGVFW 85

RESULT 11
US-09-154-750A-90
; Sequence 90, Application us/09154750A

us-09-910-582b-2.ra1

Wed Jul 16 13:41:34 2003

```

; Patent No. 6432640
; GENERAL INFORMATION:
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Polyak, Kornelia
; TITLE OF INVENTION: p53-Induced Apoptosis
; FILE REFERENCE: 1107.75357
; CURRENT APPLICATION NUMBER: US/09/154.750A
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/059,153
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/079817
; PRIOR FILING DATE: 1998-03-30
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 90
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-154-750A-90

Query Match 75.0%; Score 33; DB 4; Length 359;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGVFW 6
DB 98 GGVFW 102

RESULT 12
US-08-872-302-2
; Sequence 2, Application US/08872302
; Patent No. 5846784
; GENERAL INFORMATION:
; APPLICANT: Hitz, William D
; TITLE OF INVENTION: Fatty Acid Modifying Enzymes From
; TITLE OF INVENTION: Developing Seeds of Vernonia galamensis
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.I. duPont de Nemours and Co.
; STREET: 1007 Market Street
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: USA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/872.302
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Majarian, William R
; REGISTRATION NUMBER: P-41,173
; REFERENCE/DOCKET NUMBER: BB-1084
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-4926
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 382 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-872-302-2

Query Match 75.0%; Score 33; DB 2; Length 382;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 GGVFW 6
DB 373 GGVFW 377

RESULT 13
US-09-041-236-4
; Sequence 4, Application US/09041236
; Patent No. 6225285
; GENERAL INFORMATION:
; APPLICANT: Luo, Yuling
; APPLICANT: Xiomei, Xu
; TITLE OF INVENTION: Semaphorin K1
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,236
; FILING DATE: March 11, 1998
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: EXEL98-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 606 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-041-236-4

Query Match 75.0%; Score 33; DB 4; Length 606;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGVFW 6
DB 394 GGVFW 398

RESULT 14
US-09-134-001C-3196
; Sequence 3196, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3196
; LENGTH: 81

```



```
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3196

Query Match      72.7%; Score 32; DB 4; Length 81;
Best Local Similarity 80.0%; Pred. No. 77;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 GGVFW 6
Db      1 GGIFW 5

RESULT 15
US-09-362-831-4
; Sequence 4, Application US/09362831
; Patent No. 6306400
; GENERAL INFORMATION:
; APPLICANT: BUBLOT et al.
; TITLE OF INVENTION: AVIAN RECOMBINANT LIVE VACCINE USING, AS VECTOR, THE
; FILE OF INVENTION: AVIAN INFECTIOUS LARYNGOTRACHEITIS VIRUS
; FILE REFERENCE: 454313-2520
; CURRENT APPLICATION NUMBER: US/09/362,831
; CURRENT FILING DATE: 1999-07-28
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Infectious Laryngotracheitis Virus
US-09-362-831-4

Query Match      72.7%; Score 32; DB 4; Length 291;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGVFW 6
Db      182 GGGILW 187
```

Search completed: July 16, 2003, 13:17:05
Job time : 9.19512 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:15:24 ; Search time 13.3171 Seconds
(without alignments)
62.425 Million cell updates/sec

Title: US-09-910-582B-2
Perfect score: 44
Sequence: 1 GGGVFWQ 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues
Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
5: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pap.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	7	9	US-09-910-582B-2
2	44	100.0	7	10	US-09-782-650-1
3	37	84.1	480	10	US-09-881-752A-38
4	36	81.8	280	10	US-09-925-300-1419
5	35	79.5	416	9	US-10-260-212-2
6	35	79.5	729	10	US-09-815-242-10132
7	34	77.3	238	9	US-10-101-464A-702
8	34	77.3	346	9	US-09-917-378-8
9	34	77.3	491	9	US-10-156-761-8169
10	33	75.0	109	9	US-09-764-868-944
11	33	75.0	301	9	US-10-304-928-14
12	33	75.0	317	9	US-09-151-771-8
13	33	75.0	318	9	US-09-151-771-9
14	33	75.0	342	9	US-09-809-391-695
15	33	75.0	409	9	US-10-269-353-42
16	33	75.0	409	10	US-09-790-264-42
17	33	75.0	410	9	US-10-156-761-11310
18	33	75.0	417	10	US-09-815-242-10165
19	33	75.0	432	9	US-10-081-872-74

20	33	75.0	432	9	US-10-081-872-84	Sequence 84, Appl
21	33	75.0	432	9	US-10-081-872-86	Sequence 86, Appl
22	33	75.0	436	9	US-10-081-872-6	Sequence 6, Appl
23	33	75.0	436	9	US-10-081-872-22	Sequence 22, Appl
24	33	75.0	436	9	US-10-081-872-24	Sequence 24, Appl
25	33	75.0	436	9	US-10-081-872-28	Sequence 28, Appl
26	33	75.0	436	9	US-10-081-872-32	Sequence 32, Appl
27	33	75.0	436	9	US-10-081-872-36	Sequence 36, Appl
28	33	75.0	436	9	US-10-081-872-44	Sequence 44, Appl
29	33	75.0	436	9	US-10-081-872-46	Sequence 46, Appl
30	33	75.0	436	9	US-10-081-872-48	Sequence 48, Appl
31	33	75.0	436	9	US-10-081-872-54	Sequence 54, Appl
32	33	75.0	436	9	US-10-081-872-66	Sequence 66, Appl
33	33	75.0	437	9	US-10-269-353-30	Sequence 30, Appl
34	33	75.0	437	10	US-09-790-264-30	Sequence 30, Appl
35	33	75.0	438	10	US-09-815-242-5203	Sequence 5203, Ap
36	33	75.0	444	9	US-10-156-761-7910	Sequence 7910, Ap
37	33	75.0	450	9	US-10-269-353-34	Sequence 34, Appl
38	33	75.0	450	10	US-09-790-264-34	Sequence 34, Appl
39	33	75.0	457	9	US-10-081-872-110	Sequence 110, App
40	33	75.0	457	9	US-10-081-872-316	Sequence 316, App
41	33	75.0	460	9	US-10-081-872-314	Sequence 314, App
42	33	75.0	461	9	US-10-081-872-82	Sequence 82, Appl
43	33	75.0	464	9	US-10-081-872-126	Sequence 126, App
44	33	75.0	480	9	US-10-269-353-32	Sequence 32, Appl
45	33	75.0	480	10	US-09-790-264-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1
US-09-910-582B-2
; Sequence 2, Application US/0910582B
; Publication No. US20030045476A1
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Mackenna, Deirdre A.
; TITLE OF INVENTION: Heart Homing Conjugates
; FILE REFERENCE: P-LJ 4857
; CURRENT APPLICATION NUMBER: US/09/910, 582B
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 09/326,718
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; TYPE: PRT
; LENGTH: 7
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-09-910-582B-2

Query Match 100.0%; Score 44; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGGVFWQ 7
| | | | |
Db 1 GGGVFWQ 7

RESULT 2
US-09-782-650-1
; Sequence 1, Application US/09782650
; Patent No. US20020019350A1
; GENERAL INFORMATION:
; APPLICANT: Levine, Arnold J.
; APPLICANT: Mitterer, Artur
; APPLICANT: Falkner, Falko-Gunter
; APPLICANT: Scheifflinger, Friedrich
; APPLICANT: Dörner, Friedrich

us-09-910-582b-2.rapb

Wed Jul 16 13:41:34 2003

APPLICANT: Edwards Lifesciences Corporation

TITLE OF INVENTION: Targeted Angiogenesis
 FILE REFERENCE: 20553D-000611US
 CURRENT APPLICATION NUMBER: US/09/782,650
 CURRENT FILING DATE: 2001-02-12
 PRIOR APPLICATION NUMBER: US 09/324,079
 PRIOR FILING DATE: 1999-06-01
 PRIOR APPLICATION NUMBER: US 09/327,045
 PRIOR FILING DATE: 1999-06-07
 PRIOR APPLICATION NUMBER: PCT/US00/14988
 PRIOR FILING DATE: 2000-05-31
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 1
 LENGTH: 7
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence:targeting
 OTHER INFORMATION: molecule

US-09-782-650-1
 Query Match 100.0%; Score 44; DB 10; Length 7;
 Best Local Similarity 100.0%; Pred. No. 4.1e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVFWQ 7
 Db 1 GGGVFWQ 7

RESULT 3
 US-09-881-752A-38
 Sequence 38, Application US/09881752A
 Patent No. US20020115078A1
 GENERAL INFORMATION:
 APPLICANT: Kleanthous, Harold
 APPLICANT: Al-Garawi, Amal
 APPLICANT: Miller, Charles
 APPLICANT: Tomb, Jean-Francois
 APPLICANT: Omen, Raymond P.
 TITLE OF INVENTION: Identification of Polynucleotides
 TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides in the
 TITLE OF INVENTION: Genome
 FILE REFERENCE: 06132/041002
 CURRENT APPLICATION NUMBER: US/09/881,752A
 CURRENT FILING DATE: 2001-06-15
 PRIOR APPLICATION NUMBER: US 08/833,457
 PRIOR FILING DATE: 1997-04-01
 NUMBER OF SEQ ID NOS: 370
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 38
 LENGTH: 480
 TYPE: PRT
 ORGANISM: Helicobacter pylori
 US-09-881-752A-38

Query Match 84.1%; Score 37; DB 10; Length 480;
 Best Local Similarity 71.4%; Pred. No. 1.4e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGVFWQ 7
 Db 130 GAGIFWQ 136

RESULT 4
 US-09-925-300-1419
 Sequence 1419, Application US/09925300
 Patent No. US20020151681A1
 GENERAL INFORMATION:
 APPLICANT: Craig Rosen,
 APPLICANT: Steve Ruben

US-09-925-300-1419
 Query Match 81.8%; Score 36; DB 10; Length 280;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVFW 6
 Db 101 GGGVFW 106

RESULT 5
 US-10-260-212-2
 Sequence 2, Application US/10260212
 Publication No. US20030131380A1
 GENERAL INFORMATION:
 APPLICANT: SOCIETE DES PRODUITS NESTLE S.A.
 TITLE OF INVENTION: Coffee Mannanase
 FILE REFERENCE: 88265-6783
 CURRENT APPLICATION NUMBER: US/10/260,212
 CURRENT FILING DATE: 2002-08-27
 PRIOR APPLICATION NUMBER: PCT/EP01/01549
 PRIOR FILING DATE: 2001-02-13
 NUMBER OF SEQ ID NOS: 25
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 2
 LENGTH: 416
 TYPE: PRT
 ORGANISM: Coffea arabica
 US-10-260-212-2

Query Match 79.5%; Score 35; DB 9; Length 416;
 Best Local Similarity 71.4%; Pred. No. 2.6e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGVFWQ 7
 Db 372 GGNLFWQ 378

RESULT 6
 US-09-815-242-10132
 Sequence 10132, Application US/09815242
 Patent No. US20020061569A1
 GENERAL INFORMATION:
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Kari L.
 APPLICANT: Zyskind, Judith W.
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John D.
 APPLICANT: Carr, Grant J.
 APPLICANT: Yamamoto, Robert T.
 APPLICANT: Xu, H. Howard
 TITLE OF INVENTION: Identification of Essential Genes in
 TITLE OF INVENTION: Prokaryotes
 FILE REFERENCE: ELITRA.011A
 CURRENT APPLICATION NUMBER: US/09/815,242
 CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/191,078

US-09-815-242-10132
 Query Match 84.1%; Score 37; DB 10; Length 480;
 Best Local Similarity 71.4%; Pred. No. 1.4e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGVFWQ 7
 Db 130 GAGIFWQ 136

; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10132
 ; LENGTH: 729
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-09-815-242-10132

Query Match 79.5%; Score 35; DB 10; Length 729;
 Best Local Similarity 85.7%; Pred. No. 4.3e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGVFWQ 7
 Db 649 GGVFWQ 655
 |||||

RESULT 7
 US-10-101-464A-702
 ; Sequence 702, Application US/10101464A
 ; Publication No. US20030046728A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Strabala, Timothy
 ; APPLICANT: Nieuwenhuizen, Nicolaas
 ; APPLICANT: Higgins, Colleen M.
 ; TITLE OF INVENTION: Compositions Isolated from Plant Cells
 ; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
 ; FILE REFERENCE: 11000.1020c2
 ; CURRENT APPLICATION NUMBER: US/10/101.464A
 ; CURRENT FILING DATE: 2002-03-18
 ; PRIOR APPLICATION NUMBER: 09/704,302
 ; PRIOR FILING DATE: 2000-11-01
 ; PRIOR APPLICATION NUMBER: 09/228,986
 ; PRIOR FILING DATE: 1999-01-12
 ; PRIOR APPLICATION NUMBER: 60/162,866
 ; PRIOR FILING DATE: 1999-11-01
 ; PRIOR APPLICATION NUMBER: PCT/US00/00724
 ; PRIOR FILING DATE: 2000-01-11
 ; NUMBER OF SEQ ID NOS: 989
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 702
 ; LENGTH: 238
 ; TYPE: PRT
 ; ORGANISM: Pinus radiata
 US-10-101-464A-702

Query Match 77.3%; Score 34; DB 9; Length 238;
 Best Local Similarity 83.3%; Pred. No. 2.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGVFWQ 7
 Db 18 GGVFWQ 23
 |||||

RESULT 8
 US-09-917-378-8
 ; Sequence 8, Application US/09917378
 ; Publication No. US20030119093A1
 ; GENERAL INFORMATION:

; APPLICANT: DING, SHI-YOU
 ; APPLICANT: ADNEY, WILLIAM S.
 ; APPLICANT: VINZANT, TODD B.
 ; APPLICANT: DECKER, STEPHEN R.
 ; APPLICANT: HIMMEL, MICHAEL E.
 ; TITLE OF INVENTION: THERMAL TOLERANT MANNANASE FROM ACIDOTHERMUS
 ; TITLE OF INVENTION: CELLULOXYTICUS
 ; FILE REFERENCE: 40197.7US01
 ; CURRENT APPLICATION NUMBER: US/09/917,378
 ; CURRENT FILING DATE: 2001-07-28
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 8
 ; LENGTH: 346
 ; TYPE: PRT
 ; ORGANISM: Trichoderma reesei
 US-09-917-378-8

Query Match 77.3%; Score 34; DB 9; Length 346;
 Best Local Similarity 71.4%; Pred. No. 3.2e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGVFWQ 7
 Db 302 GGVFWQ 308
 |||||

RESULT 9
 US-10-156-761-8169
 ; Sequence 8169, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IREDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156.761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 8169
 ; LENGTH: 491
 ; TYPE: PRT
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-8169

Query Match 77.3%; Score 34; DB 9; Length 491;
 Best Local Similarity 83.3%; Pred. No. 4.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGVFW 6
 Db 283 GGVFW 288
 |||||

RESULT 10
 US-09-764-868-944
 ; Sequence 944, Application US/09764868
 ; Patent No. US20020168711A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT232
 ; CURRENT APPLICATION NUMBER: US/09/764,868
 ; CURRENT FILING DATE: 2001-01-17

Wed Jul 16 13:41:34 2003

```

; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 944
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: .SITE
; LOCATION: (21)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-944

Query Match          75.0%; Score 33; DB 9; Length 109;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGVFWQ 7
Db 84 GGGHFWR 90

RESULT 11
US-10-304-928-14
; Sequence 14, Application US/10304928
; Publication No. US20030101484A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; APPLICANT: Valpani, Nasser
; TITLE OF INVENTION: Maize Chitinases and Their Use in
; FILE REFERENCE: 1100D
; CURRENT APPLICATION NUMBER: US/10/304,928
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/522,714
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/125,915
; PRIOR FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Zea mays
US-10-304-928-14

Query Match          75.0%; Score 33; DB 9; Length 301;
Best Local Similarity 66.7%; Pred. No. 4.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVFW 6
Db 152 GGGMYW 157

RESULT 12
US-09-151-771-8
; Sequence 8, Application US/09151771
; Publication No. US20020192745A1
; GENERAL INFORMATION:
; APPLICANT: SOPHIE M. LEHAR; and
; APPLICANT: BRAYDON C. GUILD
; TITLE OF INVENTION: NOVEL APOPTOSIS GENE EI24,
; TITLE OF INVENTION: COMPOSITIONS, AND METHODS OF USE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,771
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/619,362
; FILING DATE: 21 MARCH 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, Henry N.
; REGISTRATION NUMBER: 32,073
; REFERENCE/DOCKET NUMBER: 104322.170
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 942-8400
; TELEFAX: (202) 942-8484
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,771
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/619,362
; FILING DATE: 21 MARCH 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, Henry N.
; REGISTRATION NUMBER: 32,073
; REFERENCE/DOCKET NUMBER: 104322.170
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 942-8400
; TELEFAX: (202) 942-8484
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,771
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/619,362
; FILING DATE: 21 MARCH 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, Henry N.
; REGISTRATION NUMBER: 32,073
; REFERENCE/DOCKET NUMBER: 104322.170
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 942-8400
; TELEFAX: (202) 942-8484
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-151-771-8

Query Match          75.0%; Score 33; DB 9; Length 317;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGVFW 6
Db 97 GGVFW 101

RESULT 13
US-09-151-771-9
; Sequence 9, Application US/09151771
; Publication No. US20020192745A1
; GENERAL INFORMATION:
; APPLICANT: SOPHIE M. LEHAR; and
; APPLICANT: BRAYDON C. GUILD
; TITLE OF INVENTION: NOVEL APOPTOSIS GENE EI24,
; TITLE OF INVENTION: COMPOSITIONS, AND METHODS OF USE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,771
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/619,362
; FILING DATE: 21 MARCH 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: WIXON, Henry N.
; REGISTRATION NUMBER: 32,073
; REFERENCE/DOCKET NUMBER: 104322.170
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 942-8400
; TELEFAX: (202) 942-8484
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:

```

; LENGTH: 318 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-151-771-9

Query Match 75.0%; Score 33; DB 9; Length 318;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGVFW 6
|||||
Db 98 GGVFW 102

RESULT 14

US-09-809-391-695
; Sequence 695, Application US/09809391
; Publication No. US20030049618A1
; GENERAL INFORMATION:

; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: PZ002P2
; CURRENT APPLICATION NUMBER: US/09/809,391
; CURRENT FILING DATE: 2001-03-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 761
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 695
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-809-391-695

Query Match 75.0%; Score 33; DB 9; Length 342;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGVFW 6
|||||
Db 81 GGVFW 85

RESULT 15

US-10-269-353-42
; Sequence 42, Application US/10269353
; Publication No. US2003010447A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: MPI2000-5380MNIC1M
; CURRENT APPLICATION NUMBER: US/10/269,353
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US 09/790,264
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 09/065,661
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 09/298,531
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: US 09/065,363
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 09/337,930
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/102,705
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: US 09/363,630
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 09/124,538
; PRIOR FILING DATE: 1998-07-29

; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-269-353-42

Query Match 75.0%; Score 33; DB 9; Length 409;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGVFW 6
|||||
Db 92 GGVFW 96

Search completed: July 16, 2003, 13:35:55
Job time : 14.3171 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:09:18 ; Search time 9.21951 Seconds
(without alignments)
72.991 Million cell updates/sec

Title: US-09-910-582B-2
Perfect score: 44
Sequence: 1 GGGVFWQ 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR-73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	88.6	379	2 A10825	probable 3-phenylp
2	38	86.4	411	2 D86153	hypothetical prote
3	37	84.1	465	2 G71931	probable outer mem
4	37	84.1	480	2 G64580	hypothetical prote
5	36	81.8	757	2 B75437	ABC transporter, A
6	35	79.5	131	2 AG1421	hypothetical prote
7	35	79.5	647	1 QYFEGM	phosphoenolpyruvat
8	35	79.5	713	2 JC5870	poly(beta-D-mannur
9	35	79.5	729	2 C64854	ferric-coprogen re
10	35	79.5	729	2 H90813	outer membrane rec
11	35	79.5	729	2 D85673	outer membrane rec
12	35	79.5	753	2 AE0187	probable iron-side
13	35	79.5	809	2 S32899	ferric-pseudobacti
14	34	77.3	157	2 B64245	pilin repressor pi
15	34	77.3	157	2 T33561	peptide methionine
16	34	77.3	165	2 D81412	probable peptide m
17	34	77.3	177	2 G90308	peptide methionine
18	34	77.3	177	2 G84273	peptide methionine
19	34	77.3	224	2 AE1999	ATP-binding protei
20	34	77.3	444	2 AD0338	heamin storage sys
21	34	77.3	457	2 T47005	hypothetical prote
22	34	77.3	530	2 F90893	probable kinase [i
23	34	77.3	530	2 C85724	probable kinase yd
24	34	77.3	530	2 AF0051	probable carbohydr
25	34	77.3	530	2 AE0941	probable sugar kin
26	34	77.3	530	2 B64905	sugar kinase homol
27	34	77.3	532	2 C95916	probable sugar kin
28	34	77.3	546	2 G71348	probable apolipop
29	34	77.3	900	2 T47732	probable translati

translation initia
protein C18H9.3 [1
OmpA-related prote
hypothetical prote
hypothetical prote
heme exporter prot
transport permealase
transport permealase
hypothetical prote
hypothetical prote
branched-chain ami
chitinase (EC 3.2.
probable integral
ChIB protein limpo
hypothetical prote

ALIGNMENTS

RESULT 1

AI0825

Probable 3-phenylpropionate permease STY2801 [imported] - Salmonella enterica subsp.
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C;Accession: AI0825

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se

A;Reference number: AB0502; PMID:11677608

A;Accession: AI0825

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-379 <PAR>

A;Cross-references: GB:AL513382; PIDN:CAD02757.1; PID:g16503767; GSPDB:GN00176

C;Genetics:

A;Gene: STY2801

C;Superfamily: maltose permease

Query Match 88.6% Score 39; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGVFW 6
|||||
Db 354 GGGVFW 359

RESULT 2

D86153

hypothetical protein T6A9.1 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001

C;Accession: D86153

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: D86153

A;Status: preliminary

A;Molecule type: DNA

Wed Jul 16 13:41:34 2003

us-09-910-582b-2.rpr

```

A:Residues: 1-411 <STO>
A:Cross-references: GB:AE005172; NID:99857528; PIDN:RAG00883.1; GSPDB:GN00141
C:Species: Deinococcus radiodurans
C:Genetics:
A:Map position: 1

  Query Match      86.4%; Score 38; DB 2; Length 411;
  Best Local Similarity 100.0%; Pred. No. 21;
  Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGVFWQ 7
   |||||
Db 357 GGVFWQ 362

RESULT 3
G71931
probable outer membrane protein - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 26-May-2000
C:Accession: G71931
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Werberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: G71931
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-465 <ARN>
A:Cross-references: GB:AE001478; GB:AE001439; NID:94154971; PIDN:AAD06017.1; PID:9415497
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0439
C:Superfamily: Helicobacter pylori hypothetical protein HP0209

  Query Match      84.1%; Score 37; DB 2; Length 465;
  Best Local Similarity 71.4%; Pred. No. 35;
  Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGVFWQ 7
   |||||
Db 115 GAGIFWQ 121

RESULT 4
G64580
hypothetical protein HP0487 - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 26-May-2000
C:Accession: G64580
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: G64580
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-480 <TOM>
A:Cross-references: GB:AE000563; GB:AE000511; NID:92313593; PIDN:AAD07555.1; PID:9231359
C:Superfamily: Helicobacter pylori hypothetical protein HP0209

  Query Match      84.1%; Score 37; DB 2; Length 480;
  Best Local Similarity 71.4%; Pred. No. 36;
  Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGVFWQ 7
   |||||
Db 130 GAGIFWQ 136

RESULT 5
B75437
ABC transporter, ATP-binding protein, EF-3 family - Deinococcus radiodurans (strain R
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: B75437
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567286
A:Accession: B75437
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-757 <WHI>
A:Cross-references: GB:AE001960; GB:AE000513; NID:96458833; PIDN:AAF10677.1; PID:9645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1103
A:Map position: 1

  Query Match      81.8%; Score 36; DB 2; Length 757;
  Best Local Similarity 83.3%; Pred. No. 87;
  Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGVFW 6
   |||||
Db 619 GGGFW 624

RESULT 6
AG1421
hypothetical protein lmo2776 [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AG1421
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
ok, C.; Schlueter, T.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1421
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-131 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAD00989.1; PID:g16412276; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo2776

  Query Match      79.5%; Score 35; DB 2; Length 131;
  Best Local Similarity 83.3%; Pred. No. 23;
  Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGVFW 6
   |||||
Db 67 GGGFW 72

RESULT 7
QYFGM
phosphoenolpyruvate carboxykinase (GTP) (EC 4.1.1.32) precursor, mitochondrial - frui
N:Alternate names: phosphoenolpyruvate carboxylase
C:Species: Drosophila melanogaster
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 18-Jun-1999
C:Accession: A26809
R:Gundelfinger, E.D.; Hermans-Borgmeyer, I.; Grenningloh, G.; Zopf, D.
Nucleic Acids Res. 15, 6745, 1987
```

A:Title: Nucleotide and deduced amino acid sequence of the phosphoenolpyruvate carboxykinase
A:Reference number: A26809; MUID:87316942; PMID:3114718
A:Accession: A26809
A:Molecule type: mRNA
A:Residues: 1-647 <GUN>
A:Cross-references: GB:Y00402; NID:g8326; PIDN:CAA68463.1; PID:g8327
C:Comment: This enzyme catalyzes the formation of phosphoenolpyruvate by transferring the
se activity is affected by a number of hormones regulating this metabolic process.
C:Genetics:
A:Gene: zdf4
A:Cross-references: FlyBase:FBgn0003067
C:Superfamily: phosphoenolpyruvate carboxykinase (GTP)
C:Keywords: carbon-carbon lyase; carboxy-lyase; gluconeogenesis; GTP binding; mitochondria
F:263-270/Region: nucleotide-binding motif A (P-loop) #status atypical
F:314/Active site: Cys #status predicted

Query Match 79.5%; Score 35; DB 1; Length 647;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGVFWQ 7
|||||
Db 392 GGVFWE 397

RESULT 8
JC5870
poly(beta-D-mannuronate) lyase (EC 4.2.2.3) - Pseudomonas sp.
N:Alternate names: alginate lyase 1
C:Species: Pseudomonas sp.
C:Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 20-Jun-2000
C:Accession: JC5870
R:Kraiwattanaopong, J.; Ooi, T.; Kinoshita, S.
Biosci. Biotechnol. Biochem. 61, 1853-1857, 1997
A:Title: Cloning and sequence analysis of the gene (alyII) coding for an alginate lyase
A:Reference number: JC5870; MUID:198067850; PMID:9404064
A:Accession: JC5870
A:Molecule type: DNA
A:Residues: 1-713 <KRA>
A:Cross-references: DDBJ:AB003330; NID:g2073106; PIDN:BAA19848.1; PID:g2073107
A:Experimental source: strain OS-ALG-1
C:Comment: This enzyme catalyzes the degradation of alginate by a beta-elimination mechanism
cleaved alginate oligomers.
C:Genetics:
A:Gene: alyII
C:Superfamily: Pseudomonas poly(beta-D-mannuronate) lyase
C:Keywords: carbon-oxygen lyase

Query Match 79.5%; Score 35; DB 2; Length 713;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGVFWQ 7
||:||||
Db 136 GGRIFWQ 142

RESULT 9
C64854
ferric-coprogen receptor precursor - Escherichia coli (strain K-12)
N:Alternate names: outer membrane protein fhue
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: C64854; S09262; A26875; PC4409
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: C64854
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-729 <BLAT>

A:Cross-references: GB:AE000210; GB:U00096; NID:g1787332; PIDN:AAC74186.1; PID:g17873
A:Experimental source: strain K-12, substrain MG1655
R:Sauer, M.; Hantke, K.; Braun, V.
Mol. Microbiol. 4, 427-437, 1990
A:Title: Sequence of the fhue outer-membrane receptor gene of Escherichia coli K12 an
A:Reference number: S09262; MUID:90286919; PMID:2162465
A:Accession: S09262
A:Molecule type: DNA
A:Residues: 1-362, 'C', 364-729 <SAU>
A:Cross-references: EMBL:X17615; NID:g41448; PIDN:CAA35616.1; PID:g41449
A:Experimental source: strain K-12
R:Sauer, M.; Hantke, K.; Braun, V.
J. Bacteriol. 169, 2044-2049, 1987
A:Title: Ferric-coprogen receptor FhuE of Escherichia coli: processing and sequence c
A:Reference number: A26875; MUID:87194585; PMID:3032906
A:Accession: A26875
A:Molecule type: DNA
A:Residues: 1-69 <SA2>
R: Kobayashi, Y.; Sasanuma, A.; Nishimura, A.; Kuratomi, K.
Res. Commun. Biochem. Cell Mol. Biol. 1, 157-170, 1997
A:Title: A novel 14-kilodalton protein in pl,p4-bis(5'-adenosyl)tetraphosphate (Ap4A)
A:Reference number: JC5685
A:Accession: PC4409
A:Molecule type: DNA
A:Residues: 1-47 <KOB>
C:Genetics:
A:Gene: fhuE
A:Map position: 16 min
C:Function:
A:Description: required for uptake of iron(III) via coprogen, ferrioxamine B and rhod
C:Superfamily: ferrichrome-iron receptor; tonB-dependent receptor amino-terminal homo
C:Keywords: iron transport; membrane protein; receptor
F:1-36/Domain: signal sequence #status predicted <SIG>
F:37-729/Product: outer membrane protein fhuE #status predicted <WAT>
F:97-230/Domain: tonB-dependent receptor amino-terminal homology <TNN>
F:462-729/Domain: tonB-dependent receptor carboxyl-terminal homology <TNC>

Query Match 79.5%; Score 35; DB 2; Length 729;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGVFWQ 7
|||||
Db 649 GGVNMQ 655

RESULT 10
H90813
outer membrane receptor for ferric iron uptake ECs1480 [imported] - Escherichia coli
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: H90813
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: H90813
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-729 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA034903.1; PID:g13360944; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs1480
C:Superfamily: ferric-pseudobactin receptor; tonB-dependent receptor amino-terminal h

Query Match 79.5%; Score 35; DB 2; Length 729;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGVFWQ 7
|||||

Wed Jul 16 13:41:34 2003

us-09-910-582b-2.rpr

```

Db          649 GGGVNWQ 655

RESULT 11
outer membrane receptor for ferric iron uptake [imported] - Escherichia coli (strain O157:H7)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: D85673
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, M.W.;
Miller, L.; Grotbeck, E.J.; Davis, A.; Dimalanta, E.; Potamousis, K.; Apodaca, N.;
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85673
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-729 <STO>
A:Cross-references: GB:AE005174; NID:q12514649; PIDN:AAG55848.1; GSPDB:GN00145; UWGP:217
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: fhuE
C:Superfamily: ferric-pseudobactin receptor; tonB-dependent receptor amino-terminal homology
Query Match          79.5%; Score 35; DB 2; Length 729;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY          1 GGGVFWQ 7
          |||||
Db          649 GGGVNWQ 655

RESULT 12
probable iron-siderophore receptor YPO1537 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C:Accession: AE0187
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, T.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AE0187
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-753 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC90360.1; PID:gl5979580; GSPDB:GN00175
C:Genetics:
A:Gene: YPO1537
C:Superfamily: ferric-pseudobactin receptor; tonB-dependent receptor amino-terminal homology
Query Match          79.5%; Score 35; DB 2; Length 753;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY          1 GGGVFWQ 7
          |||||
Db          671 GGGVNWQ 677

RESULT 13
ferric-pseudobactin receptor pupB precursor - Pseudomonas putida
C:Species: Pseudomonas putida
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 29-Sep-1999
C:Accession: S32899
R:Koster, M.; van de Vossen, J.; Leong, J.; Weisbeek, P.J.
Mol. Microbiol. 8, 591-601, 1993
A:Title: Identification and characterization of the pupB gene encoding an inducible ferric
A:Reference number: S32899; MUID:93316856; PMID:8392140
A:Accession: S32899
A:Molecule type: DNA
A:Residues: 1-809 <KOS>
A:Cross-references: GB:X73598; GB:S63481; NID:g403024; PIDN:CAA51995.1; PID:g581467
C:Genetics:
A:Gene: pupB
C:Superfamily: ferric-pseudobactin receptor; tonB-dependent receptor amino-terminal homology
Query Match          79.5%; Score 35; DB 2; Length 809;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY          1 GGGVFWQ 7
          |||||
Db          735 GGGVNWQ 741

RESULT 14
piliin repressor pilB homolog MG408 - Mycoplasma genitalium
C:Species: Mycoplasma genitalium
C:Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 07-Dec-1999
C:Accession: B64245
R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.;
M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick,
C.A.; Venter, J.C.
Science 270, 397-403, 1995
A:Title: The minimal gene complement of Mycoplasma genitalium.
A:Reference number: A64200; MUID:96026346; PMID:7569993
A:Accession: B64245
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-157 <TIGR>
A:Cross-references: GB:U039726; GB:LA3967; NID:gl0461113; PID:gl046122; TIGR:MG408
A:Experimental source: strain G-37
C:Genetics:
A:Gene: ggcFW
C:Superfamily: peptide methionine sulfoxide reductase
Query Match          77.3%; Score 34; DB 2; Length 157;
Best Local Similarity 83.3%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY          1 GGGVFW 6
          |||||
Db          7 GGGCFW 12

RESULT 15
peptide methionine sulfoxide reductase pmsR - Mycoplasma pneumoniae (strain ATCC 2934)
C:Species: Mycoplasma pneumoniae
A:Variety: ATCC 29342
C:Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999
C:Accession: S73561
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkkl, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
A:Reference number: S73327; MUID:97105885; PMID:8948633
A:Accession: S73561
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-157 <HIM>
A:Cross-references: EMBL:AE000023; GB:U000089; NID:gl673893; PIDN:AA895883.1; PID:gl67
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C:Genetics:

```

A:Gene: pmsR
A:Genetic code: SGC3
C:Superfamily: peptide methionine sulfoxide reductase

Query Match 77.3% Score 34; DB 2; Length 157;
Best Local Similarity 83.3%; Pred. NO. 42;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGVFW 6
||| ||
Db 7 GGGCFW 12

Search completed: July 16, 2003, 13:16:10
Job time : 10.2195 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:04:48 ; Search time 4.78049 Seconds
(without alignments)
60.733 Million cell updates/sec

Title: US-09-910-582B-2
Perfect score: 44
Sequence: 1 GGGVFWQ 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	36	81.8	275	1 RR40_HUMAN	Q9nqt5 homo sapien
2	35	79.5	647	1 PPCK_DROME	P20007 drosophila
3	35	79.5	729	1 FHUE_ECOLI	P16869 escherichia
4	35	79.5	809	1 PUPB_PSEPU	P38047 pseudomonas
5	34	77.3	157	1 MSRA_MYCGE	P47648 mycoplasma
6	34	77.3	157	1 MSRA_MYCPN	P75188 mycoplasma
7	34	77.3	155	1 MSRA_CAMJE	Q9phn0 campylobact
8	34	77.3	177	1 MSRA_HALNI	Q9hgg0 halobacteri
9	34	77.3	177	1 MSRA_SULSO	Q97y45 sulfolobus
10	34	77.3	191	1 MSRA_RALSO	Q8ylc6 ralstonia s
11	34	77.3	530	1 YDEV_ECOLI	P77432 escherichia
12	34	77.3	546	1 LNT_TREPA	O83279 treponema p
13	34	77.3	900	1 IF3B_ARATH	O49160 arabidopsis
14	34	77.3	1061	1 OAR_MYXXA	P38370 myxococcus
15	33	75.0	163	1 YE99_MYCPN	P75288 mycoplasma
16	33	75.0	319	1 CBIB_SALTY	Q05600 salmonella
17	33	75.0	347	1 Y576_METJA	O57996 methanococc
18	33	75.0	358	1 PIG8_MOUSE	O61070 mus musculu
19	33	75.0	359	1 PIG8_HUMAN	O14681 homo sapien
20	33	75.0	417	1 YNFW_ECOLI	P43531 escherichia
21	33	75.0	427	1 CH11_COCIM	P54196 coccidioid
22	33	75.0	458	1 HRA2_MOUSE	Q9j1v5 mus musculu
23	33	75.0	1355	1 ATC3_YEAST	P39524 saccharomyc
24	33	75.0	1440	1 POLG_JAEVN	P14403 j genome po
25	33	75.0	1780	1 POLG_MWEV	P05769 m genome po
26	33	75.0	3432	1 POLG_JAEV1	P27395 j genome po
27	33	75.0	3432	1 POLG_JAEV5	P19110 j genome po
28	33	75.0	3432	1 POLG_JAEVJ	P32886 j genome po
29	32	72.7	238	1 Y819_CHLTR	O84826 chlamydia t
30	32	72.7	292	1 LICB_HAEIN	P14182 haemophilus
31	32	72.7	348	1 POTD_ECOLI	P23861 escherichia
32	32	72.7	348	1 POTD_SALTY	P97012 salmonella
33	32	72.7	375	1 YY10_MYCLE	Q49721 mycobacteri

34	32	72.7	378	1 YF69_AQUAE	O67513 aquifex aeo
35	32	72.7	384	1 HYMA_EMENI	O60032 emericella
36	32	72.7	423	1 CH14_TRIHA	P48827 trichoderma
37	32	72.7	431	1 PHOR_ECOLI	P08400 escherichia
38	32	72.7	449	1 YKE6_YEAST	P36091 saccharomyc
39	32	72.7	458	1 YM77_YEAST	Q05031 saccharomyc
40	32	72.7	484	1 GUNA_XANCP	P19487 xanthomonas
41	32	72.7	521	1 LAG3_MOUSE	Q61790 mus musculu
42	32	72.7	622	1 PPCC_CHICK	P05153 gallus gall
43	32	72.7	622	1 PPCC_HUMAN	P35558 homo sapien
44	32	72.7	622	1 PPCC_MOUSE	Q922v4 mus musculu
45	32	72.7	622	1 PPCC_RAT	P07379 rattus norv

ALIGNMENTS

RESULT 1
RR40_HUMAN
ID RR40_HUMAN STANDARD; PRT; 275 AA.
AC Q9NQT5: Q9Y3A8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Exosome complex exonuclease RRP40 (EC 3.1.13.-) (Ribosomal RNA
DE processing protein 40) (p10) (CGI-102).
GN RRP40.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=21125821; PubMed=1110791;
RA Brouwer R., Allmang C., Rajmakers R., van Aarsen Y., Egberts W.V.,
RA Petfalski E., van Venrooij W.J., Tollervey D., Pruijn G.J.M.;
RT "Three novel components of the human exosome.";
RL J. Biol. Chem. 276:6177-6184(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin, and Uterus;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBDJ databases.
RN [3]
RP SEQUENCE OF 56-275 FROM N.A.
RX MEDLINE=20272150; PubMed=10810093;
RA Lai C.-H., Chou C.-Y., Chang L.-Y., Liu C.-S., Lin W.-C.;
RT "Identification of novel human genes evolutionarily conserved in
RL Caenorhabditis elegans by comparative proteomics.";
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=99396719; PubMed=10465791;
RA Allmang C., Petfalski E., Podtelejnikov A., Mann M., Tollervey D.,
RA Mitchell P.;
RT "The yeast exosome and human PM-Scl are related complexes of 3'-5'
RL exonucleases.";
RN [5]
RP Genes Dev. 13:2148-2158(1999).
CC -1- FUNCTION: COMPONENT OF THE EXOSOME 3->5 EXORIBONUCLEASE COMPLEX.
CC REQUIRED FOR THE 3' PROCESSING OF THE 7S PRE-RNA TO THE MATURE
CC 5.8S RNA (BY SIMILARITY).
CC -1- SUBUNIT: COMPONENT OF THE EXOSOME MULTISUBUNIT RIBONUCLEASE COMPLEX
CC COMPOSED OF AT LEAST 11 PROTEINS: RRP40, RRP41/SKI16, RRP42,
CC RRP43, RRP44/DIS3, PM/SCL-75, RRP46, CSL4 AND PM/SCL-100 (ONLY IN
CC THE NUCLEAR COMPLEX). ALSO ASSOCIATED WITH THE GTPASE RAN (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR; NUCLEOLUS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

Wed Jul 16 13:41:35 2003

us-09-910-582b-2.rsp

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AF281132; AAF82133.1; -
 CC EMBL; BC002437; AAH02437.1; -
 DR EMBL; BC008880; AAH08880.1; -
 DR EMBL; AF151860; AAD34097.1; -
 DR MIM; 606489; -
 KW Exosome; Hydrolyase; Nuclease; Exonuclease; rRNA processing;
 KW Nuclear protein; RNA-binding. NARACSRVRV -> MLERARGCAF (IN REF. 3).
 FT CONFLICT 56 65 S -> G (IN REF. 3).
 FT CONFLICT 95 95
 SQ SEQUENCE 275 AA; 29572 MW; 26432214A199166 CRC64;

Query Match 81.8%; Score 36; DB 1; Length 275;
 Best Local Similarity 83.3%; Pred. No. 14;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGVFW 6
 Db 96 GGGVFW 101
 |||||:1

RESULT 2
 PPCK_DROME STANDARD; PRT; 647 AA.
 ID PPCK_DROME STANDARD; PRT; 647 AA.
 AC P20007; Q9V8J0;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phosphoenolpyruvate carboxylase [GTP] (EC 4.1.1.32)
 DE (Phosphoenolpyruvate carboxylase) (PEPCK).
 GN PEPCK OR ZDF4 OR CGL7725.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OK NCBI_TaxID=7227;
 [1]
 SEQUENCE FROM N.A.
 RP MEDLINE=87316942; PubMed=3114719;
 RX Gundelfinger E.D., Hermans-Borgmeyer I., Grenningloh G., Zopf D.;
 RT "Nucleotide and deduced amino acid sequence of the
 RT phosphoenolpyruvate carboxylase (GTP) from *Drosophila*
 RT melanogaster".
 RL Nucleic Acids Res. 15:6745-6745(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Bessie P.V., Bereman B.P., Bhandari D., Brottier P.,
 RA Borkova D., Botchan M.R., Boulc J., Brokstein P., Center A., Chandra I.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Cadenot L.B., Davies P.,
 RA Cherry J.M., Cowley S., Dahlke C., Davenport L., Desmet G., Dunn P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Fierres S., Fleischmann W.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Glasser K.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iqbalwan C.,
 RA Jalali M., Kalush K.A., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhou X., Zhu S., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zheng L.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 CC -1- CATALYTIC ACTIVITY: GTP + oxaloacetate = GDP + phosphoenolpyruvate
 CC + CO(2).
 CC -1- PATHWAY: RATE-LIMITING GLUCONEOGENIC ENZYME WHOSE ACTIVITY IS
 CC AFFECTED BY A NUMBER OF HORMONES REGULATING THIS METABOLIC
 CC PROCESS.
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOENOLPYRUVATE CARBOXYKINASE [GTP]
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; Y00402; CAA68463.1; -
 EMBL; AE003799; AAF57676.1; -
 DR PIR; A26809; QYFFGM.
 DR FlyBase; FBgn0003067; PEPCK.
 DR InterPro; IPR000364; PEP_carboxykin.
 DR Pfam; PF00821; PEPCK; 1.
 DR ProDom; PD004738; PEP_carboxykin; 1.
 DR PROSITE; PS00505; PEPCK_GTP; 1.
 DR Gluconeogenesis; Lyase; Decarboxylase; GTP-binding.
 NP_BIND 263 270 GTP (POTENTIAL).
 FT ACT_SITE 314 314 BY SIMILARITY.
 FT CONFLICT 302 302 E -> V (IN REF. 1).
 FT CONFLICT 408 408 O -> R (IN REF. 1).
 SQ SEQUENCE 647 AA; 71129 MW; 0DB81B3D9E1B1FBB CRC64;

Query Match 79.5%; Score 35; DB 1; Length 647;
 Best Local Similarity 83.3%; Pred. No. 47;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GGVFWQ 7
 Db 392 GGVFW 397
 |||||:

RESULT 3
 FHUE_ECOLI STANDARD; PRT; 729 AA.
 ID FHUE_ECOLI STANDARD; PRT; 729 AA.
 AC P16869; P77292;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE FhuE receptor precursor (Outer-membrane receptor for Fe(III)-coprogen,
 DE Fe(III)-ferrioxamine B and Fe(III)-rhodotriacetic acid).
 GN FHUE OR B1102.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OC NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=90286919; PubMed=2162465;

RA Sauer U., Hantke K., Braun V.;
 RT "Sequence of the fhuE outer-membrane receptor gene of Escherichia
 RL coli K12 and properties of mutants."; Mol. Microbiol. 4:427-437(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RL "The complete genome sequence of Escherichia coli K-12."; Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97061202; PubMed=8905232;
 RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
 RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Iano M., Horiuchi T.;
 RL "A 718-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 12.7-28.0 min region on the linkage map."; DNA Res. 3:137-155(1996).
 RN [4]
 RP SEQUENCE OF 37-50.
 RC STRAIN=K12;
 RX MEDLINE=87194585; PubMed=3032906;
 RA Sauer M., Hantke K., Braun V.;
 RT "Ferric-coprogen receptor FhuE of Escherichia coli: processing and
 RT sequence common to all TonB-dependent outer membrane receptor
 RT proteins."; J. Bacteriol. 169:2044-2049(1987).
 RL J. Bacteriol. 169:2044-2049(1987).
 CC -1- FUNCTION: REQUIRED FOR THE UPTAKE OF IRON(III) VIA COPROGEN,
 CC FERRIOXAMINE B, AND RHODOTOLIC ACID.
 CC -1- SUBCELLULAR LOCATION: Outer membrane.
 CC -1- INDUCTION: FOR INDUCTION THE TONB AND THE EXBB PROTEIN HAVE TO
 CC BE ACTIVE.
 CC -1- SIMILARITY: LOCAL TO OTHER TONB-DEPENDENT RECEPTOR PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X17615; CAA35616.1; -
 CC DR EMBL: AE000210; AAC74186.1; -
 CC DR EMBL: D90745; BAA35909.1; -
 CC DR EMBL: D90746; BAA35917.1; -
 CC DR PIR: S09262; S09262.
 CC DR PIR: A26875; A26875.
 CC DR EcoGene: EG10306; fhuE.
 DR InterPro: IPR000531; TonB_boxC.
 DR Pfam: PF00593; TonB_boxC; 1.
 DR PROSITE: PS00430; TONB_DEPENDENT_REC_1; 1.
 DR PROSITE: PS01156; TONB_DEPENDENT_REC_2; 1.
 KW Outer membrane; Iron transport; Transport; TonB box; Signal; Receptor;
 KW Complete proteome.
 FT SIGNAL 1 36
 FT CHAIN 37 729 FHU E RECEPTOR.
 FT SITE 42 49 TONB BOX.
 FT SITE 712 729 TONB C-TERMINAL BOX.
 FT MUTAGEN 44 44 V->P: ABOLISHES TRANSPORT ACTIVITY.
 FT MUTAGEN 46 46 V->P: ABOLISHES TRANSPORT ACTIVITY.
 FT MUTAGEN 163 163 G->A,V,N,D,H: IMPAIRED TRANSPORT AND GROWTH.
 FT CONFLICT 363 363 G -> C (IN REF. 1).

SQ SEQUENCE 729 AA; 81232 MW; 68ACEE7D110F76CD CRC64;
 Query Match 79.5%; Score 35; DB 1; Length 729;
 Best Local Similarity 85.7%; Pred. No. 53;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGVFWQ 7
 IIII II
 Db 649 GGGVNWQ 655
 RESULT 4
 PUPB_PSEPU
 ID PUPB_PSEPU STANDARD; PRT; 809 AA.
 AC P38047;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Ferric-pseudobactin BN7/BN8 receptor precursor.
 GN PUPB.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WCS358;
 RX MEDLINE=93316856; PubMed=8392140;
 RA Koster M., van de Vossen J., Leong J., Weisbeek P.J.;
 RT "Identification and characterization of the pnpB gene encoding an
 RT inducible ferric-pseudobactin receptor of Pseudomonas putida
 RT WCS358."; Mol. Microbiol. 8:591-601(1993).
 RL Mol. Microbiol. 8:591-601(1993).
 CC -1- FUNCTION: SPECIFIC RECEPTOR FOR THE SIDEROPHORES FERRIC
 CC PSEUDOBACTIN BN8 AND FERRIC PSEUDOBACTIN BN7, IRON CHELATING
 CC MOLECULES THAT ALLOW THE ORGANISM TO EXTRACT IRON FROM THE
 CC ENVIRONMENT, SPECIALLY UNDER IRON-RESTRICTED CONDITIONS.
 CC -1- SUBCELLULAR LOCATION: Outer membrane.
 CC -1- INDUCTION: BY IRON LIMITATION IN ADDITION TO THE PRESENCE OF ONE
 CC OF THE TWO COGNATE PSEUDOBACTINS BN8 OR BN7.
 CC -1- SIMILARITY: LOCAL TO OTHER TONB-DEPENDENT RECEPTOR PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X73598; CAA51995.1; -
 CC DR PIR: S32899; S32899.
 DR InterPro: IPR000531; TonB_boxC.
 DR Pfam: PF00593; TonB_boxC; 1.
 DR PROSITE: PS00430; TONB_DEPENDENT_REC_1; FALSE_NEG.
 DR PROSITE: PS01156; TONB_DEPENDENT_REC_2; 1.
 KW Outer membrane; Iron transport; Transport; Signal; Receptor.
 FT SIGNAL 1 45 POTENTIAL.
 FT CHAIN 46 809 FERRIC-PSEUDOBACTIN BN7/BN8 RECEPTOR.
 FT DOMAIN 156 160 POLY-SER.
 FT SITE 792 809 TONB C-TERMINAL BOX.
 SQ SEQUENCE 809 AA; 88389 MW; 0B339F6E788A8C0D CRC64;
 Query Match 79.5%; Score 35; DB 1; Length 809;
 Best Local Similarity 85.7%; Pred. No. 58;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGVFWQ 7
 IIII II
 Db 735 GGGVNWQ 741
 RESULT 5

RESULT 6
MSRA MYCPN

RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
 RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
 RA Jags K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
 RA Whitehead S., Barrett B.G.,
 RT "The genome sequence of the food-borne pathogen *Campylobacter jejuni*
 RT reveals hypervariable sequences.";
 RL Nature 403:665-668(2000).
 CC -1- FUNCTION: Has an important function as a repair enzyme for
 CC proteins that have been inactivated by oxidation. Catalyzes the
 CC reversible oxidation-reduction of methionine sulfoxide in proteins
 CC to methionine (By similarity).
 CC -1- CATALYTIC ACTIVITY: Protein L-methionine + oxidized thiorodoxin =
 CC protein L-methionine S-oxide + reduced thiorodoxin.
 CC -1- SIMILARITY: BELONGS TO THE MSRA MET SULFOXIDE REDUCTASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AL139075; CAB75273.1;
 CC HSSP; P54149; lFVG.
 CC InterPro; IPR002569; PMSR.
 CC Pfam; PF01625; PMSR; 1.
 CC ProDom; PD003489; PMSR; 1.
 CC TIGRFAMS; TIGR00401; msrA; 1.
 CC Oxidoreductase; Complete proteome.
 CC ACT_SITE 10 10 BY SIMILARITY.
 CC SEQUENCE 165 AA; 18896 MW; 8DF4366859BF4F13 CRC64;
 CC -----
 CC Query Match 77.3%; Score 34; DB 1; Length 165;
 CC Best Local Similarity 83.3%; Pred. No. 20;
 CC Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 1 GGGVFW 6
 CC III II
 CC Db 7 GGGCFW 12
 CC
 CC RESULT 8
 CC MSRA_HALN1 STANDARD; PRT; 177 AA.
 CC AC QHQGO;
 CC DT 15-JUN-2002 (Rel. 41, Created)
 CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Peptide methionine sulfoxide reductase msrA (EC 1.8.4.6) (Protein-
 CC methionine-S-oxide reductase) (Peptide Met(O) reductase).
 CC GN MSRA OR VNG180G.
 CC OS Halobacterium sp. (strain NRC-1).
 CC OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 CC OC Halobacteriaceae; Halobacterium.
 CC OX NCBI_TaxID=54091;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=20504483; PubMed=11016950;
 CC RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
 RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Irsenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
 RT "Genome sequence of Halobacterium species NRC-1".
 CC Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 CC -1- FUNCTION: Has an important function as a repair enzyme for
 CC proteins that have been inactivated by oxidation. Catalyzes the
 CC reversible oxidation-reduction of methionine sulfoxide in proteins

CC to methionine (By similarity).
 CC -1- CATALYTIC ACTIVITY: Protein L-methionine + oxidized thiorodoxin =
 CC protein L-methionine S-oxide + reduced thiorodoxin.
 CC -1- SIMILARITY: BELONGS TO THE MSRA MET SULFOXIDE REDUCTASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE005045; AAG19555.1;
 CC HSSP; P54149; lFVG.
 CC InterPro; IPR002569; PMSR.
 CC Pfam; PF01625; PMSR; 1.
 CC ProDom; PD003489; PMSR; 1.
 CC TIGRFAMS; TIGR00401; msrA; 1.
 CC Oxidoreductase; Complete proteome.
 CC ACT_SITE 12 12 BY SIMILARITY.
 CC SEQUENCE 177 AA; 19202 MW; 337FBD2E1CFDD8EC CRC64;
 CC -----
 CC Query Match 77.3%; Score 34; DB 1; Length 177;
 CC Best Local Similarity 83.3%; Pred. No. 21;
 CC Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 1 GGGVFW 6
 CC III II
 CC Db 9 GGGCFW 14
 CC
 CC RESULT 9
 CC MSRA_SULSO STANDARD; PRT; 177 AA.
 CC AC Q9YI45;
 CC DT 15-JUN-2002 (Rel. 41, Created)
 CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Peptide methionine sulfoxide reductase msrA (EC 1.8.4.6) (Protein-
 CC methionine-S-oxide reductase) (Peptide Met(O) reductase).
 CC GN MSRA OR MSR OR SSO1503.
 CC OS Sulfolobus solfataricus.
 CC OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 CC OC Sulfolobus.
 CC OX NCBI_TaxID=2287;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 CC RX MEDLINE=21332296; PubMed=11427726;
 CC RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Aways M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duquet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sengen C.W., Van der Oost J.;
 RT "The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2.";
 CC Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 CC -1- FUNCTION: Has an important function as a repair enzyme for
 CC proteins that have been inactivated by oxidation. Catalyzes the
 CC reversible oxidation-reduction of methionine sulfoxide in proteins
 CC to methionine (By similarity).
 CC -1- CATALYTIC ACTIVITY: Protein L-methionine + oxidized thiorodoxin =
 CC protein L-methionine S-oxide + reduced thiorodoxin.
 CC -1- SIMILARITY: BELONGS TO THE MSRA MET SULFOXIDE REDUCTASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE006765; AAK41726.1; -
DR InterPro: IPR002569; PMSR.
DR Pfam; PF01625; PMSR; 1.
DR ProDom; PD003489; PMSR; 1.
DR TIGRFAMs; TIGR00401; msrA; 1.
KW Oxidoreductase; Complete proteome.
FT ACT_SITE 10 BY SIMILARITY.
SQ SEQUENCE 177 AA: 20570 MW; 02EEFLAC5CA458B6 CRC64;

Query Match 77.3%; Score 34; DB 1; Length 177;
Best Local Similarity 83.3%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGVFW 6
Db 7 GGGCFW 12

RESULT 10
MSRA_RALSO STANDARD; PRT; 191 AA.
ID MSRA_RALSO
AC Q8YIC6;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Peptide methionine sulfoxide reductase msrA (EC 1.8.4.6) (Protein-
methionine-S-oxide reductase) (peptide Met(O) reductase).
GN MSRA OR RSC0764 OR RSC05089.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
Arlat M., Billault A., Brothier P., Camus J.C., Cattolico L.,
Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,
Weissenbach J., Boucher C.A.;
RA "Genome sequence of the plant pathogen Ralstonia solanacearum";
RL Nature 415:497-502(2002).
CC -!- FUNCTION: Has an important function as a repair enzyme for
proteins that have been inactivated by oxidation. Catalyzes the
reversible oxidation-reduction of methionine sulfoxide in proteins
to methionine (By similarity).
CC -!- CATALYTIC ACTIVITY: Protein L-methionine + oxidized thioredoxin =
protein L-methionine S-oxide + reduced thioredoxin.
CC -!- SIMILARITY: BELONGS TO THE MSRA MET SULFOXIDE REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL646060; CAD14294.1; -
DR InterPro: IPR002569; PMSR.
DR TIGRFAMs; TIGR00401; msrA; 1.
KW Oxidoreductase; Complete proteome.
FT ACT_SITE 21 BY SIMILARITY.
SQ SEQUENCE 191 AA: 21023 MW; D8CCAB398FA27535 CRC64;

Query Match 77.3%; Score 34; DB 1; Length 191;
Best Local Similarity 83.3%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGVFW 6
Db 7 GGGCFW 12

RESULT 11
YDEV_ECOLI STANDARD; PRT; 530 AA.
ID YDEV_ECOLI
AC P77432; Q99894;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical sugar kinase ydev.
GN YDEV OR B1511.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=K12; PubMed=9097039;
RX MEDLINE=97251357; PubMed=8649811;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
Motomura K., Nakade S., Nakamura Y., Nashimoto H., Nishio Y.,
Oshima T., Saito N., Sampei G., Seki Y., Sivasubram S.,
Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
Yamamoto Y., Horiuchi T.;
RA "A 570-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map";
RL DNA Res. 3:363-377(1996).
RN [3]
RP SEQUENCE OF 182-495 FROM N.A.
RX MEDLINE=96243037; PubMed=8649811;
RA Das R., Reddy E.P., Chatterjee D., Andrews D.W.;
RA "Identification of a novel Bcl-2 related gene, BRAG-1, in human
glioma";
RL Oncogene 12:947-951(1996).
CC -!- SIMILARITY: BELONGS TO THE FUCOKINASE / GLUCOKINASE /
GLYCEROKINASE / XYLULOKINASE FAMILY.
CC -!- CAUTION: WAS THOUGHT BY REF. 3 TO BE A HUMAN SEQUENCE AND WAS
CALLED BY THEM BRAG1 (BRAIN-RELATED APOPTOSIS GENE) (BRAG-1)
WITH A ROLE IN APOPTOSIS. THE DNA SEQUENCE OF THE REGION THEY
SEQUENCED IS MORE THAN 99% IDENTICAL TO THAT OF THIS E.COLI
GENE. FURTHERMORE THEY CLAIM 'EXTENSIVE SIMILARITY TO THE
BCL-2 FAMILY OF GENES.' SUCH A SIMILARITY IS NOT SIGNIFICANT
AND THIS PROTEIN IS MUCH MORE LIKELY TO BE A SUGAR KINASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000248; AAC74584.1; -
DR EMBL; D90793; BAA15191.1; -
DR EMBL; D90794; BAA15198.1; -
DR EMBL; S82185; AAC17184.1; -
DR EcoGene; EGI3804; ydev.
DR InterPro: IPR000577; FGGY_kin.

DR Pfam; PF00370; FGGY; 1.
 DR Pfam; PF02782; FGGY C; 1.
 DR PROSITE; PS00933; FGGY_KINASES_1; FALSE_NEG.
 DR PROSITE; PS00445; FGGY_KINASES_2; FALSE_NEG.
 KW Hypothetical protein; transferase; kinase; Complete proteome.
 FT CONFLICT 490 495 PDPEKH -> TRPGA (IN REF. 2).
 SQ SEQUENCE 530 AA; 57544 MW; CBC3B1E7C8982063 CRC64;

Query Match 77.3%; Score 34; DB 1; Length 530;
 Best Local Similarity 83.3%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGVFWQ 7
 II III
 Db 273 GGTFWQ 278

RESULT 12

LNT_TREPA STANDARD; PRT; 546 AA.
 ID LNT_TREPA
 AC O83279;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Apolipoprotein N-acyltransferase (EC 2.3.1.-) (ALP N-acyltransferase).
 GN LNT OR TP0252.
 OS Treponema pallidum.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
 OX NCBI_TaxID=160;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nichols;
 RX MEDLINE=98332770; PubMed=9665876;
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
 RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
 RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
 RA Kralak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
 RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
 RA Venter J.C.;
 RT "Complete genome sequence of Treponema pallidum, the syphilis
 spirochete.";
 RL Science 281:375-388(1998).
 CC -1- FUNCTION: TRANSFERS THE FATTY ACYL GROUP ON MEMBRANE LIPOPROTEINS
 (BY SIMILARITY).
 CC -1- PATHWAY: Lipoproteins biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE APOLIPOPROTEIN N-ACYLTRANSFERASE
 FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE001206; AAC65237.1; -
 DR TIGR; TP0252; -
 DR InterPro; IPR003010; Ntlse/CNhydrtse.
 DR Pfam; PF00795; CN_Hydrolase; 1.
 DR Transferase; Acyltransferase; Transmembrane; Complete proteome.
 FT TRANSMEM 14 34 POTENTIAL.
 FT TRANSMEM 62 82 POTENTIAL.
 FT TRANSMEM 85 105 POTENTIAL.
 FT TRANSMEM 122 142 POTENTIAL.
 FT TRANSMEM 151 171 POTENTIAL.
 FT TRANSMEM 194 214 POTENTIAL.
 FT TRANSMEM 490 510 POTENTIAL.
 FT TRANSMEM 514 534 POTENTIAL.
 SQ SEQUENCE 546 AA; 61513 MW; 06E8041A3FB8821E CRC64;

Query Match 77.3%; Score 34; DB 1; Length 546;
 Best Local Similarity 83.3%; Pred. No. 61;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGVFWQ 7
 II III
 Db 108 GGAFWQ 113

RESULT 13

IF38_ARATH STANDARD; PRT; 900 AA.
 ID IF38_ARATH
 AC O49160;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable eukaryotic translation initiation factor 3 subunit 8 (eIF3
 DE p110) (p105).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99065153; PubMed=9849901;
 RA Karniol B., Yahalom A., Kwok S., Tsuge T., Matsui M., Deng X.-W.,
 RA Chamovitz D.A.;
 RT "The Arabidopsis homologue of an eIF3 complex subunit associates with
 RT the COP9 complex.";
 RL FEBS Lett. 439:173-179(1998).
 CC -1- FUNCTION: BINDS TO THE 40S RIBOSOME AND PROMOTES THE BINDING OF
 CC METHIONYL-TRNAI AND MRNA (BY SIMILARITY).
 CC -1- SUBUNIT: EIF-3 IS COMPOSED OF AT LEAST 10 DIFFERENT SUBUNITS (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE EIF38 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF040102; AAC83464.1; -
 DR InterPro; IPR000717; PCI.
 DR Pfam; PF01399; PCI; 1.
 DR SMART; SM00088; PINT; 1.
 DR KW Initiation factor; Protein biosynthesis.
 SQ SEQUENCE 900 AA; 102989 MW; 3C03FEDBF59AB5D1 CRC64;

Query Match 77.3%; Score 34; DB 1; Length 900;
 Best Local Similarity 71.4%; Pred. No. 98;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGVFWQ 7
 III :II
 Db 829 GGGGVWQ 835

RESULT 14

OAR_MYXXA STANDARD; PRT; 1061 AA.
 ID OAR_MYXXA
 AC P38370;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE OAR protein precursor.
 GN OAR.
 OS Myxococcus xanthus.
 OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
 OC Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.

Wed Jul 16 13:41:35 2003

Query Match 75.0%; Score 33; DB 1; Length 163;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GGVFW 6
Db 21 GGVFW 25

Search completed: July 16, 2003, 13:13:11
Job time : 6.78049 secs

OX NCBL_TaxID-34;
RN SEQUENCE FROM N.A., AND SEQUENCE OF 776-796.
RC STRAIN-DZF1;
RX MEDLINE-93328680; PubMed-8335633;
RA Martinez-Canamero M., Munoz-Dorado J., Farez-Vidal E., Inouye M.,
RA Inouye S.;
RT "Oar, a 115-kilodalton membrane protein required for development of
RT Myxococcus xanthus";
RL J. Bacteriol. 175:4756-4763(1993).
CC -!- FUNCTION: REQUIRED FOR CELLULAR ADHESION DURING FRUITING BODY
CC FORMATION.
CC -!- SUBCELLULAR LOCATION: OUTER-MEMBRANE ASSOCIATED.
CC -!- SIMILARITY: LOCAL, TO TONB-DEPENDENT RECEPTOR PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S64103; AAB27614.1;
DR PIR; A40609; A40609.
KW Outer membrane; Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 1061 OAR PROTEIN.
FT SEQUENCE 1061 AA; 114455 MW; EA8C077296352EF0 CRC64;
SQ

Query Match 77.3%; Score 34; DB 1; Length 1061;
Best Local Similarity 85.7%; Pred. No. 11e-02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGVFWQ 7
Db 561 GGVFWQ 567
|||||:1

RESULT 15
YE99_MYCPN STANDARD; PRT; 163 AA.
AC P75288;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MPN499 (P02_orf163).
GN MPN499 OR MP344.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBL_TaxID-2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 29342 / M129;
RX MEDLINE-97105885; PubMed-8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF000032; AAB95991.1;
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 163 AA; 19095 MW; 10CE2D2F077DFEAC CRC64;
SQ

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:07:18 ; Search time 18.9512 Seconds
(without alignments)
76.107 Million cell updates/sec

Title: US-09-910-582b-2
Perfect score: 44
Sequence: 1 GGGVFWQ 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues 671580
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_21.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phase.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriaph.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	93.2	395	10 Q8RVL3	Q8rvl3 lycopersico
2	41	93.2	399	10 Q93WT4	Q93wt4 lycopersico
3	39	88.6	279	2 Q9F7Z2	Q9f7z2 rhizobium 1
4	39	88.6	379	16 Q9EZ24	Q9ez24 salmoneilla
5	39	88.6	1245	13 Q9YGH8	Q9ygh8 scophthalmu
6	38	86.4	301	5 Q9NLG8	Q9nlg8 leishmania
7	38	86.4	328	4 Q8WY77	Q8wy77 homo sapien
8	38	86.4	411	10 Q9FZ29	Q9fz29 arabidopsis
9	37	84.1	465	16 Q9ZLY6	Q9zly6 helicobacte
10	37	84.1	480	16 Q25Z31	Q25z31 helicobacte
11	36	81.8	79	4 Q9NYS3	Q9nys3 homo sapien
12	36	81.8	213	16 Q9RCX6	Q9rcx6 streptomyce
13	36	81.8	243	11 Q9CWX2	Q9cwx2 mus musculu
14	36	81.8	757	16 Q9RVC6	Q9rv62 deinococcus
15	35	79.5	92	5 Q95UE4	Q95ue4 platyphora
16	35	79.5	98	5 Q95UD7	Q95ud7 platyphora

17	35	79.5	109	5 Q95UE1	Q95ue1 platyphora
18	35	79.5	110	5 Q95UE5	Q95ue5 platyphora
19	35	79.5	111	5 Q95UD1	Q95ud1 platyphora
20	35	79.5	115	5 Q95UD5	Q95ud5 platyphora
21	35	79.5	117	5 Q95UE6	Q95ue6 platyphora
22	35	79.5	119	5 Q95UD3	Q95ud3 platyphora
23	35	79.5	119	5 Q95UD2	Q95ud2 platyphora
24	35	79.5	120	5 Q95UE3	Q95ue3 platyphora
25	35	79.5	122	5 Q9GS65	Q9gs65 xylocopa ir
26	35	79.5	123	5 Q95UD4	Q95ud4 platyphora
27	35	79.5	123	5 Q95UC9	Q95uc9 platyphora
28	35	79.5	124	5 Q9GS49	Q9gs49 xylocopa fr
29	35	79.5	125	5 Q95UD0	Q95ud0 platyphora
30	35	79.5	125	5 Q95UC6	Q95uc6 desmogramma
31	35	79.5	126	5 Q9GS53	Q9gs53 xylocopa mi
32	35	79.5	126	5 Q95UC7	Q95uc7 platyphora
33	35	79.5	130	5 Q9GS61	Q9gs61 xylocopa pu
34	35	79.5	131	16 Q8Y3Q7	Q8y3q7 listeria mo
35	35	79.5	139	5 Q9GS73	Q9gs73 xylocopa tr
36	35	79.5	140	5 Q9GS60	Q9gs60 xylocopa sc
37	35	79.5	140	5 Q9GS59	Q9gs59 xylocopa n1
38	35	79.5	140	5 Q9GS51	Q9gs51 xylocopa gu
39	35	79.5	142	5 Q9GS68	Q9gs68 xylocopa er
40	35	79.5	143	5 Q9GS66	Q9gs66 xylocopa vi
41	35	79.5	144	5 Q9GS50	Q9gs50 xylocopa tr
42	35	79.5	146	5 Q9GS58	Q9gs58 xylocopa fl
43	35	79.5	152	5 Q9GS54	Q9gs54 xylocopa ap
44	35	79.5	153	5 Q9GS75	Q9gs75 brevinea
45	35	79.5	153	5 Q9GS74	Q9gs74 xylocopa ml

ALIGNMENTS

RESULT 1

Q8RVL3 ID Q8RVL3 PRELIMINARY; PRT; 395 AA.
AC Q8RVL3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Inactive endo-beta-mannanase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;

Q8RVL3 ID Q8RVL3 PRELIMINARY; PRT; 395 AA.
AC Q8RVL3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Inactive endo-beta-mannanase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;

Q8RVL3 ID Q8RVL3 PRELIMINARY; PRT; 395 AA.
AC Q8RVL3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Inactive endo-beta-mannanase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;

Q8RVL3 ID Q8RVL3 PRELIMINARY; PRT; 395 AA.
AC Q8RVL3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Inactive endo-beta-mannanase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;

Q8RVL3 ID Q8RVL3 PRELIMINARY; PRT; 395 AA.
AC Q8RVL3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Inactive endo-beta-mannanase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;

Q8RVL3 ID Q8RVL3 PRELIMINARY; PRT; 395 AA.
AC Q8RVL3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Inactive endo-beta-mannanase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;

Q8RVL3 ID Q8RVL3 PRELIMINARY; PRT; 395 AA.
AC Q8RVL3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Inactive endo-beta-mannanase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;

Q8RVL3 ID Q8RVL3 PRELIMINARY; PRT; 395 AA.
AC Q8RVL3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Inactive endo-beta-mannanase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;

Q8RVL3 ID Q8RVL3 PRELIMINARY; PRT; 395 AA.
AC Q8RVL3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Inactive endo-beta-mannanase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;

Q8RVL3 ID Q8RVL3 PRELIMINARY; PRT; 395 AA.
AC Q8RVL3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Inactive endo-beta-mannanase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;

Q8RVL3 ID Q8RVL3 PRELIMINARY; PRT; 395 AA.
AC Q8RVL3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Inactive endo-beta-mannanase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;

Q8RVL3 ID Q8RVL3 PRELIMINARY; PRT; 395 AA.
AC Q8RVL3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Inactive endo-beta-mannanase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;

Q8RVL3 ID Q8RVL3 PRELIMINARY; PRT; 395 AA.
AC Q8RVL3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Inactive endo-beta-mannanase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;

Q8RVL3 ID Q8RVL3 PRELIMINARY; PRT; 395 AA.
AC Q8RVL3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Inactive endo-beta-mannanase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;

Q8RVL3 ID Q8RVL3 PRELIMINARY; PRT; 395 AA.
AC Q8RVL3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Inactive endo-beta-mannanase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;

Q8RVL3 ID Q8RVL3 PRELIMINARY; PRT; 395 AA.
AC Q8RVL3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Inactive endo-beta-mannanase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;

Q8RVL3 ID Q8RVL3 PRELIMINARY; PRT; 395 AA.
AC Q8RVL3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Inactive endo-beta-mannanase.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_taxid=4081;

Wed Jul 16 13:41:35 2003

```

DE Endo-beta-mannanase (EC 3.2.1.78).
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. TRUST; TISSUE=RIPENING FRUIT PERICARP;
RA Bourgault R., Bewley J.D.;
RT "A cDNA encoding an endo-beta-mannanase expressed in ripening tomato
RL fruit of the cultivar Trust.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY046588; AAK57760.1; -
DR InterPro: IPR001547; GH_5;
DR Pfam: PF00150; cellulase; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 399 AA; 45338 MW; 805A7D7866E5FF6B CRC64;

Query Match 93.2%; Score 41; DB 10; Length 399;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVFWQ 7
DB 355 GGGLEWQ 361

RESULT 3
O9F722 PRELIMINARY; PRT; 279 AA.
AC O9F722;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Rhizobium-adhering protein RapB.
CN RAPB
OS Rhizobium leguminosarum (biovar trifolii).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=386;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R200;
RX MEDLINE=21135131; PubMed=11238962;
RA Ausmees N., Jacobsson K., Lindberg M.;
RT "A unipolarly located, cell-surface-associated agglutinin, RapA,
RT belongs to a family of Rhizobium-adhering proteins (Rap) in Rhizobium
RT leguminosarum bv. trifolii.";
RL Microbiology 147:549-559(2001).
DR EMBL: AF265223; AAG18520.1; -
SQ SEQUENCE 279 AA; 30283 MW; AC9BEA71BD010A1D CRC64;

Query Match 88.6%; Score 39; DB 2; Length 279;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVFW 6
DB 177 GGGVFW 182

RESULT 4
O9E224 PRELIMINARY; PRT; 379 AA.
AC O9E224;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE PnuD (Putative MFS family transport protein) (Putative
DE 3-phenylpropionate permease).
DN PNUD OR HCAT OR STM2554 OR STY2801.
OS Salmonella typhimurium, and

OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S. typhimurium;
RA Jeong H., Zhu N., Roth J.R.;
RT "Improved assimilation of nicotinamide mononucleotide by alteration of
RT a novel MFS transporter in Salmonella enterica.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S. typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen K., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=S. typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feitwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
DR EMBL: AF306512; AAG42315.1; -
DR EMBL: AE008816; AAL21448.1; -
DR EMBL: AL627275; CAD02757.1; -
DR InterPro: IPR000576; Lacy_symp.
DR Pfam: PF01306; Lacy_symp; 1.
DR PRINTS; PR00174; LACYSMPORT.
DR Hypothetical protein; Complete proteome.
FT VARIANT 2 A -> V.
FT SEQUENCE 379 AA; 41456 MW; 34E4AD5F4C3101C0 CRC64;

Query Match 88.6%; Score 39; DB 16; Length 379;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGVFW 6
DB 354 GGGVFW 359

RESULT 5
O9YGH8 PRELIMINARY; PRT; 1245 AA.
AC O9YGH8;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Insulin receptor (Fragment).
OS Scopthalmus maximus (turbot).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostei;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Scopthalmidae; Scopthalmus.
OX NCBI_TaxID=52904;
RN [1]
RP SEQUENCE FROM N.A.

```


RA Elies G., Duval H., Bonnet G., Wolff J., Boeuf G., Boujard D.;
 RT "Tubulin and insulin-like growth factor-1 receptors: cDNAs
 cloning and messenger RNAs expression during development.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN
 CC TYROSINE PHOSPHATE.

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 CC EMBL; AJ224994; CAA12279.1; -

DR HSP; P06213; IIRK.

DR InterPro: IPR000494; EGFR_L_domain.

DR InterPro: IPR000719; Euk_pkinase.

DR InterPro: IPR003961; FN_III.

DR InterPro: IPR002174; Furin-like.

DR InterPro: IPR002011; RTK_kinaseII.

DR InterPro: IPR001245; Tyr_pkinase.

DR Pfam: PF00041; fn3; 2.

DR Pfam: PF00757; Furin-like; 1.

DR Pfam: PF00069; pkinase; 1.

DR Pfam: PF01030; Recep_L_domain; 2.

DR PRINTS; PR00109; TYRKINASE.

DR ProDom: PD000001; Euk_pkinase; 1.

DR SMART; SM00060; FN3; 2.

DR SMART; SM00261; FU; 1.

DR SMART; SM00219; TYRK; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

DR PROSITE; PS00235; RECEPTOR_TYR_KIN_II; 1.

KW Glycoprotein; Phosphorylation; Receptor; Transmembrane;

KW Tyrosine-protein kinase.

FT NON_TER 1 1

FT CHAIN 626 1245

FT CHAIN 626 1245 INSULIN RECEPTOR, BETA SUBUNIT.

FT CHAIN 626 1245 INSULIN RECEPTOR, BETA SUBUNIT.

FT CHAIN 626 1245 D3DDC5A0C7104C25 CRC64;

SQ SEQUENCE 1245 AA; 139917 MW; D3DDC5A0C7104C25 CRC64;

Query Match 88.0%; Score 39; DB 13; Length 1245;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVFW 6

Db 839 GGGVFW 844

|||||

RESULT 6

Q9NLG8

ID Q9NLG8 PRELIMINARY; PRT; 301 AA.

AC Q9NLG8;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Probable mdj6 (Fragment).

GN LM26.474.

OS Leishmania major.

OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.

OX NCBI_TaxID=5664;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=FRIEDLIN;

RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL160493; CAB98092.1; -

DR HSP; P08622; LBQZ.

DR InterPro: IPR001623; DnaJ_N.

DR Pfam; PF00226; DnaJ; 1.

DR SMART; SM00271; DnaJ; 1.

DR PROSITE; PS50076; DnaJ_2; 1.

FT NON_TER 1

SQ SEQUENCE 301 AA; 33246 MW; 39A67B7A57315892 CRC64;

Query Match 86.4%; Score 38; DB 5; Length 301;

Best Local Similarity 83.3%; Pred. No. 55;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGVFW 6

Db 93 GGGVFW 98

|||||

RESULT 7

Q8WY77

ID Q8WY77 PRELIMINARY; PRT; 328 AA.

AC Q8WY77;

DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE PF3999.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Qin W.X., Zhou X.M., Zhang P.P., Jiang H.Q., Huang Y., Wan D.F.,

RA Gu J.R.;

RT "Novel human cDNA clones with function of inhibiting cancer cell

RT growth.";

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF258571; AAG23774.1; -

SQ SEQUENCE 328 AA; 35127 MW; 23D5FD43B6E2D72E CRC64;

Query Match 86.4%; Score 38; DB 4; Length 328;

Best Local Similarity 85.7%; Pred. No. 61;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGVFWQ 7

Db 259 GGVFWQ 265

|||||

RESULT 8

Q9FZ29

ID Q9FZ29 PRELIMINARY; PRT; 411 AA.

AC Q9FZ29;

DT 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE T6A9.1 protein (1-4).

GN T6A9.1 OR ATIG02310.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,

RA Altafi H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,

RA Buehler E., Chao Q., Chiu J., Choi E., Gonzalez A.,

RA Hwang B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,

RA Lenz C., Liu A., Liu S., Mukharsy N., Pham P., Sakano H., Shinn P.,

RA Toriumi M., Vaysberg M., Yu G., Ecker J., Theologis A., Davis R.W.;

RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,

RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,

RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,

RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,

RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,

RA Ecker J., Theologis A., Davis R.W.;

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC064879; AAG00883.1; -

DR EMBL; AY081352; AAL91241.1; -

DR InterPro: IPR001547; GH_5.

DR Pfam; PF00150; cellulase; 1.

Wed Jul 16 13:41:35 2003

```
KW Hydrolase.
SQ SEQUENCE 411 AA; 46290 MW; B72ECB54C4D7F218 CRC64;

Query Match
Best Local Similarity 86.4%; Score 38; DB 10; Length 411;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGVFWQ 7
DB 357 GGVFWQ 362

RESULT 9
Q9ZLY6 PRELIMINARY; PRT; 465 AA.
AC Q9ZLY6;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Putative outer membrane protein.
GN JHP0439.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RA "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
DR EMBL; AE001478; AAD06017.1;
DR InterPro; IPR003678; HP_OMP_2.
DR Pfam; PF02521; HP_OMP_2; 1.
KW Complete proteome.
SQ SEQUENCE 465 AA; 53480 MW; A2182C32E94AC1E4 CRC64;

Query Match
Best Local Similarity 84.1%; Score 37; DB 16; Length 465;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGVFWQ 7
DB 115 GAGIFWQ 121

RESULT 10
O25231 PRELIMINARY; PRT; 480 AA.
AC O25231;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein HP0487.
GN HP0487.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
```

```
Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
Venter J.C.;
RA "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547(1997).
DR EMBL; AE000563; AAD07555.1;
DR TIGR; HP0487;
DR InterPro; IPR003678; HP_OMP_2.
DR Pfam; PF02521; HP_OMP_2; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 480 AA; 55176 MW; 35DBDB0C7E74CF7 CRC64;

Query Match
Best Local Similarity 84.1%; Score 37; DB 16; Length 480;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGVFWQ 7
DB 130 GAGIFWQ 136

RESULT 11
Q9NYS3 PRELIMINARY; PRT; 79 AA.
AC Q9NYS3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Apoptosis-related protein PNAS-3 (Fragment).
GN PNAS-3 OR BA3J10.7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX Yu W.-Q., Sun B.-Z., Zhao Z.-L., Chai Y.-B., Li Z., Zhu F., Lu F.,
RA Yan W., Yang H.;
RA "PNAS-3, human acute promyelocytic leukemia cell line NB4's apoptosis-
RT related gene.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Laird G.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF229833; AAF42918.1;
DR EMBL; AL138752; CAC2319.1;
FT NON_TER 1
SQ SEQUENCE 79 AA; 8443 MW; 105E6C4FCBC80980 CRC64;

Query Match
Best Local Similarity 81.8%; Score 36; DB 4; Length 79;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGVFWQ 6
DB 11 GGVFWQ 16

RESULT 12
Q9RCX6 PRELIMINARY; PRT; 213 AA.
AC Q9RCX6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative integral membrane protein.
GN SCO0934 OR SCMI0.22C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomyces.
```

```

OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., James K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S., Taylor K.,
RA Seger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RL coelicolor A3(2).";
DR EMBL: AL133469; CAB63183.1;
SQ SEQUENCE 213 AA; 21467 MW; 112488011A7D7592 CRC64;

Query Match 81.8%; Score 36; DB 16; Length 213;
Best Local Similarity 83.3%; Pred. No. 86;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGVFW 6
Db 199 GGGVFW 204

RESULT 13
Q9CVM2 PRELIMINARY; PRT; 243 AA.
AC Q9CVM2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 2310005D06R1k protein (fragment).
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikolaio I., Pesole G., Quackenbush J.,
RA Schirln L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Bojelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK006281; BAB24501.1;
DR MGD: MGI:1913612; 2310005D06R1k.
FT NON_TER 1
SQ SEQUENCE 243 AA; 26372 MW; 5D1F643409DB80F6 CRC64;

Query Match 81.8%; Score 36; DB 11; Length 243;
Best Local Similarity 83.3%; Pred. No. 99;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 GGGVFW 6
Db 64 GGGVFW 69

RESULT 14
Q9RVC6 PRELIMINARY; PRT; 757 AA.
AC Q9RVC6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ABC transporter, ATP-binding protein, EF-3 family.
GN DR1103.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RL radiodurans R1.";
DR EMBL: AE001960; AAF10677.1;
DR TIGR: DR1103;
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transportr.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; ABC_transportr; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_2.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 757 AA; 82415 MW; 8B4BF092CA56F50E CRC64;

Query Match 81.8%; Score 36; DB 16; Length 757;
Best Local Similarity 83.3%; Pred. No. 3.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGVFW 6
Db 619 GGGVFW 624

RESULT 15
Q95UE4 PRELIMINARY; PRT; 92 AA.
AC Q95UE4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Phosphoenolpyruvate carboxykinase (Fragment).
GN PEPCK.
OS Platyphora opima.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Phytophaga; Chrysomeloidea; Chrysomelidae;
OC Chrysomelinae; Platyphora.
OX NCBI_TaxID=172874;
RN [1]
RP SEQUENCE FROM N.A.
RA Termonia A., Pasteels J.M., Windsor D.M., Milinkovitch M.C.;
RT "Dual Chemical Sequestration, a Key Mechanism in Transitions Among
Ecological Specializations.";

```

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY055492; AAL11339.1; -
DR InterPro; IPR000364; PEP_carboxykin.
DR Pfam; PF00821; PEPCK; 1.
DR PROSITE; PS00505; PEPCK_GTP; UNKNOWN_1.
KW Kinase; Pyruvate.

FT NON_TER 1
FT NON_TER 92 92
SQ SEQUENCE 92 AA; 9753 MW; D41C121A310F74F8 CRC64;

Query Match 79.5%; Score 35; DB 5; Length 92;
Best Local Similarity 83.3%; Pred. No. 52;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGVEFWQ 7
Db 86 GGVEFE 91
|||||

Search completed: July 16, 2003, 13:15:09
Job time : 20.9512 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:04:13 ; Search time 23.9024 Seconds
(without alignments)
39.023 Million cell updates/sec

Title: US-09-910-582B-3

Perfect score: 43

Sequence: 1 HGRVRPH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	43	100.0	7	22	Peptide which sele
2	43	100.0	7	22	Heart homing pepti
3	36	83.7	31	22	Human secreted pro
4	35	81.4	48	22	Human immune/haema
5	35	81.4	63	22	Propionibacterium
6	35	81.4	90	22	Human digestive sy
7	35	81.4	90	22	Human liver associ
8	35	81.4	90	23	Human liver antige
9	35	81.4	201	22	Human prostate can
10	35	81.4	210	22	Human prostate can

11	35	81.4	210	22	Human prostate can
12	35	81.4	256	22	Propionibacterium
13	35	81.4	266	22	Human gastric canc
14	35	81.4	1024	23	Herbicidally activ
15	34	79.1	203	22	Novel human diagno
16	34	79.1	277	18	Penicillium decumb
17	34	79.1	494	22	Drosophila melanog
18	34	79.1	605	16	WD-40 domain-contg
19	34	79.1	1156	21	Feline foamy virus
20	33	76.7	100	20	Human breast tumou
21	33	76.7	224	22	Propionibacterium
22	33	76.7	225	22	T. indica Cox1 int
23	33	76.7	704	22	Novel human diagno
24	33	76.7	704	22	Novel human diagno
25	32	74.4	19	21	Phosphatidic acid
26	32	74.4	54	21	Human secreted pro
27	32	74.4	70	22	Propionibacterium
28	32	74.4	83	22	Propionibacterium
29	32	74.4	119	21	Zea mays protein f
30	32	74.4	120	22	Propionibacterium
31	32	74.4	130	22	Propionibacterium
32	32	74.4	130	22	Novel human diagno
33	32	74.4	140	21	A. vitis hypersens
34	32	74.4	140	22	Novel human diagno
35	32	74.4	142	22	Novel human diagno
36	32	74.4	142	22	Propionibacterium
37	32	74.4	155	21	Zea mays protein f
38	32	74.4	165	22	Novel human diagno
39	32	74.4	196	22	Gene #26 human sec
40	32	74.4	225	22	Novel human diagno
41	32	74.4	246	22	Drosophila melanog
42	32	74.4	284	19	Human phosphatidic
43	32	74.4	285	19	Human phosphatidic
44	32	74.4	285	20	Human phosphatase
45	32	74.4	297	22	C. glutamicum SRT

ALIGNMENTS

RESULT 1

AAB30896
ID AAB30896 standard; peptide: 7 AA.

XX AC AAB30896;

XX DT 02-APR-2001 (first entry)

XX DE Peptide which selectively binds to normal cardiac endothelium.

XX DE Cardiac endothelium; angiogenic factor; vascular endothelium;

XX KW peripheral vascular disease; cardiovascular disease; angiogenesis;

XX KW cardiac neovascularisation.

XX OS Unidentified.

XX OS XX

XX PN WO200075329-A1.

XX PD 14-DEC-2000.

XX PF 31-MAY-2000; 2000WO-US14988.

XX PR 07-JUN-1999; 99US-0327045.

XX PR (EDWA-) EDWARDS LIFESCIENCES CORP.

XX PA (BAXT) BAXTER AG.

XX PI Levine AJ, Mitterer A, Falkner F, Scheiflinger F, Dorner F;

XX DR WPI; 2001-091212/10.

XX PT New chimeric molecules having an angiogenic factor linked to a

XX PT targeting molecule that binds to a vascular endothelium, useful for

Wed Jul 16 13:41:35 2003

PT increasing cardiac neovascularisation, or treating peripheral vascular
PT and cardiovascular diseases -
PS Disclosure; Page 27; 67pp; English.
XX
XX AAB30895-99 represent targeting molecules, which are used to produce the
CC chimeric molecules of the invention. AAB30895-98 selectively bind to
CC normal cardiac endothelium. The specification describes a chimeric
CC molecule comprising an angiogenic factor linked to a targeting molecule
CC that specifically binds to a vascular endothelium. The chimeric
CC molecules are useful for treatment of peripheral vascular or
CC cardiovascular diseases. Specifically, they are useful for inducing or
CC inhibiting angiogenesis, for increasing cardiac neovascularisation in
CC ischemic tissue in the peripheral vascular system.
XX
XX
SQ Sequence 7 AA;
Query Match. 100.0%; Score 43; DB 22; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HGRVRPH 7
DB 1 HGRVRPH 7
RESULT 2
AAB50798
ID AAB50798 standard; peptide; 7 AA.
XX
XX AAB50798;
XX
XX 21-MAR-2001 (first entry)
XX
XX Heart homing peptide SEQ ID NO: 3.
XX
XX Heart homing peptide; cardiovascular disease; ischaemic disease;
XX gene therapy.
XX
XX Synthetic.
XX
XX WO200075174-A1.
XX
XX 14-DEC-2000.
XX
XX 31-MAY-2000; 2000WO-US15088.
XX
XX 07-JUN-1999; 99US-0326718.
XX
XX (BURN-) BURNHAM INST.
XX
XX Ruoslahti E, Mackenna DA;
XX
XX WPI; 2001-071059/08.
XX
XX Novel heart homing peptide that selectively homes to normal ischaemic
XX and cardiac tissue useful for targeting ischaemic tissues for treating
XX ischaemic and cardiovascular diseases such as atherosclerosis and
XX stenosis
XX
XX Claim 2; Page 55; 70pp; English.
XX
XX The present invention provides a number of heart homing peptides which
XX selectively home to cardiac tissue. These can be used in the treatment of
XX cardiovascular and ischaemic diseases, such as atherosclerosis,
XX thrombosis, restenosis, vasculitis, atherosclerotic aneurysms, myocardial
XX hypertrophy, congenital heart diseases, ischaemic heart disease and
XX anginas, acquired valvular/endocardial diseases, primary myocardial
XX diseases, cardiac tumours and arrhythmias.
XX
XX
SQ Sequence 7 AA;
Query Match 100.0%; Score 43; DB 22; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HGRVRPH 7
DB 1 HGRVRPH 7
RESULT 3
AAB01660
ID AAE01660 standard; peptide; 31 AA.
XX
XX AAE01660;
XX
XX 17-JUL-2001 (first entry)
XX
XX Human secreted protein fragment, SEQ ID NO:214.
XX
XX Human; secreted protein; proliferative disorder; cancer; tumour;
XX foetal abnormality; developmental abnormality; haematopoietic disorder;
XX immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
XX inflammation; allergy; neurological disorder; Alzheimer's disease;
XX Parkinson's disease; cognitive disorder; schizophrenia; asthma;
XX skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
XX cardiovascular disorder; angiogenic disorder; kidney disorder;
XX gastrointestinal disorder; pregnancy-related disorder;
XX endocrine disorder; infection; wound healing; vulnery;
XX cell culture; chemotaxis; food additive; gene therapy;
XX binding partner identification.
XX
XX Homo sapiens.
XX
XX OS
XX Key Location/Qualifiers
XX Misc-difference 7 /note= "Xaa is any of the naturally occurring
XX /L-amino acids"
XX
XX Misc-difference 22 /note= "Xaa is any of the naturally occurring
XX /L-amino acids"
XX
XX Misc-difference 25 /note= "Xaa is any of the naturally occurring
XX /L-amino acids"
XX
XX Misc-difference 26 /note= "Xaa is any of the naturally occurring
XX /L-amino acids"
XX
XX Misc-difference 29 /note= "Xaa is any of the naturally occurring
XX /L-amino acids"
XX
XX Misc-difference 30 /note= "Xaa is any of the naturally occurring
XX /L-amino acids"
XX
XX WO200134623-A1.
XX
XX
XX 17-MAY-2001.
XX
XX 01-NOV-2000; 2000WO-US30037.
XX
XX 05-NOV-1999; 99US-0163577.
XX
XX 30-JUN-2000; 2000US-0215137.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Komatsoulis GA, Moore PA;
XX
XX WPI; 2001-316490/33.
XX
XX Nucleic acids encoding 29 human secreted polypeptides, useful for
XX preventing, diagnosing and/or treating e.g. cancers, Parkinson's
XX disease and diabetic retinopathy -
XX
XX Disclosure; Page 530; 535pp; English.
XX
XX PS
XX

CC AAD05389-AAD05473 represent cDNAs corresponding to 29 human secreted
 CC protein genes, and AAE01546-AAE01630 represent the proteins they encode.
 CC AAE01631-AAE01660 represent human secreted protein fragments or variants.
 CC The secreted proteins and their genes are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 29 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiogenic disorders, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein fragment referred to in the disclosure of the invention.

XX SQ Sequence 31 AA;

Query Match 83.7%; Score 36; DB 22; Length 31;

Best Local Similarity 85.7%; Pred. No. 2.9;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 HGRVRPH 7

Db 11 HGRVPPH 17

RESULT 4

AA082605
 ID AA082605 standard; Protein; 48 AA.

AC AA082605;

XX DT 07-NOV-2001 (first entry)

XX DE Human immune/haematopoietic antigen SEQ ID NO:10198.

XX KW Human; Immune; haematopoietic; immune/haematopoietic antigen; cancer;
 XX cytotstatic; gene therapy; vaccine; metastasis.

XX OS Homo sapiens.

XX PN WC0200157182-A2.

XX PD 09-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01354.

XX XX 31-JAN-2000; 2000US-0179065.

XX PR 04-FEB-2000; 2000US-0180628.

XX PR 24-FEB-2000; 2000US-0184664.

XX PR 02-MAR-2000; 2000US-0186350.

XX PR 16-MAR-2000; 2000US-0189874.

XX PR 17-MAR-2000; 2000US-0190076.

XX PR 18-APR-2000; 2000US-0198123.

XX PR 19-MAY-2000; 2000US-0205515.

XX PR 07-JUN-2000; 2000US-0209467.

XX PR 28-JUN-2000; 2000US-0214886.

XX PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235836.
 PR 27-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0239335.
 PR 13-OCT-2000; 2000US-0239337.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.

that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/haematopoietic-related diseases, especially cancers and cancer metastases of haematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/haematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAK82169 represent sequences used in the exemplification of the present invention.

XX Sequence 48 AA;

Query Match 81.4%; Score 35; DB 22; Length 48;

Best Local Similarity 71.4%; Pred. No. 7.2;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGRVRPH 7

DB 34 HGELRPH 40

RESULT 5

AAU59994 standard; Protein; 63 AA.

XX AC AAU59994;

XX DT 27-FEB-2002 (first entry)

XX DE Propionibacterium acnes immunogenic protein #20890.

XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis; uveitis; endophthalmitis; bone; joint; central nervous system; ELISA; inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay; dermatological; osteopathic; neuroprotectant.

XX OS Propionibacterium acnes.

XX PN WO200181581-A2.

XX PD 01-NOV-2001.

XX PF 20-APR-2001; 2001WO-US12865.

XX PR 21-APR-2000; 2000US-199047P.

XX PR 02-JUN-2000; 2000US-208841P.

XX PR 07-JUL-2000; 2000US-216747P.

XX PA (CORI-) CORIXA CORP.

XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX DR WPI; 2001-616774/71.

XX DR N-PSDB; AAS59607.

XX PT Propionibacterium acnes polypeptides and nucleic acids useful for vaccinating against and diagnosing infections, especially useful for treating acne vulgaris.

XX PS Example 1; SEQ ID NO 21189; 1069pp; English.

XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic polypeptides. The proteins and their associated DNA sequences are used in the treatment, prevention and diagnosis of medical conditions caused by P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention.

20-OCT-2000; 2000US-0241808.
20-OCT-2000; 2000US-0241809.
20-OCT-2000; 2000US-0241826.
01-NOV-2000; 2000US-0244617.
08-NOV-2000; 2000US-0246474.
08-NOV-2000; 2000US-0246475.
08-NOV-2000; 2000US-0246476.
08-NOV-2000; 2000US-0246477.
08-NOV-2000; 2000US-0246478.
08-NOV-2000; 2000US-0246523.
08-NOV-2000; 2000US-0246524.
08-NOV-2000; 2000US-0246525.
08-NOV-2000; 2000US-0246526.
08-NOV-2000; 2000US-0246527.
08-NOV-2000; 2000US-0246528.
08-NOV-2000; 2000US-0246532.
08-NOV-2000; 2000US-0246609.
08-NOV-2000; 2000US-0246610.
08-NOV-2000; 2000US-0246611.
17-NOV-2000; 2000US-0246613.
17-NOV-2000; 2000US-0249207.
17-NOV-2000; 2000US-0249208.
17-NOV-2000; 2000US-0249209.
17-NOV-2000; 2000US-0249210.
17-NOV-2000; 2000US-0249211.
17-NOV-2000; 2000US-0249212.
17-NOV-2000; 2000US-0249213.
17-NOV-2000; 2000US-0249214.
17-NOV-2000; 2000US-0249215.
17-NOV-2000; 2000US-0249216.
17-NOV-2000; 2000US-0249217.
17-NOV-2000; 2000US-0249218.
17-NOV-2000; 2000US-0249219.
17-NOV-2000; 2000US-0249244.
17-NOV-2000; 2000US-0249245.
17-NOV-2000; 2000US-0249264.
17-NOV-2000; 2000US-0249265.
17-NOV-2000; 2000US-0249297.
17-NOV-2000; 2000US-0249299.
17-NOV-2000; 2000US-0249300.
17-NOV-2000; 2000US-0250160.
01-DEC-2000; 2000US-0250391.
05-DEC-2000; 2000US-0251030.
05-DEC-2000; 2000US-0251988.
05-DEC-2000; 2000US-0256719.
06-DEC-2000; 2000US-0251479.
08-DEC-2000; 2000US-0251856.
08-DEC-2000; 2000US-0251868.
08-DEC-2000; 2000US-0251869.
08-DEC-2000; 2000US-0251989.
08-DEC-2000; 2000US-0251990.
11-DEC-2000; 2000US-0254097.
05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

N-PSDB; AAK5386.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis.

Claim 11; SEQ ID NO 10198; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome.

CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 63 AA;

Query Match 81.4%; Score 35; DB 22; Length 63;

Best Local Similarity 100.0%; Pred. No. 9.7;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 GRRVPH 7

|||||

Db 56 GRRVPH 61

RESULT 6

AAM91924

ID AAM91924 standard; Protein; 90 AA.

XX AC AAM91924;

XX DT 06-NOV-2001 (first entry)

XX DE Human digestive system antigen SEQ ID NO: 1273.

KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum.

XX OS Homo sapiens.

XX PN WO200155314-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01324.

XX PR 31-JAN-2000; 2000US-0179065.

XX PR 04-FEB-2000; 2000US-0180628.

XX PR 24-FEB-2000; 2000US-0184664.

XX PR 02-MAR-2000; 2000US-0186350.

XX PR 16-MAR-2000; 2000US-0189874.

XX PR 17-MAR-2000; 2000US-0190076.

XX PR 18-APR-2000; 2000US-0198123.

XX PR 19-MAY-2000; 2000US-0205515.

XX PR 07-JUN-2000; 2000US-0209467.

XX PR 28-JUN-2000; 2000US-0214886.

XX PR 30-JUN-2000; 2000US-0215135.

XX PR 07-JUL-2000; 2000US-0216647.

XX PR 07-JUL-2000; 2000US-0216880.

XX PR 11-JUL-2000; 2000US-0217487.

XX PR 11-JUL-2000; 2000US-0217496.

XX PR 14-JUL-2000; 2000US-0218290.

XX PR 26-JUL-2000; 2000US-0220963.

XX PR 26-JUL-2000; 2000US-0220964.

XX PR 14-AUG-2000; 2000US-0224518.

XX PR 14-AUG-2000; 2000US-0224519.

XX PR 14-AUG-2000; 2000US-0225213.

XX PR 14-AUG-2000; 2000US-0225214.

XX PR 14-AUG-2000; 2000US-0225266.

XX PR 14-AUG-2000; 2000US-0225267.

XX PR 14-AUG-2000; 2000US-0225268.

XX PR 14-AUG-2000; 2000US-0225270.

XX PR 14-AUG-2000; 2000US-0225447.

XX PR 14-AUG-2000; 2000US-0225757.

XX PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231988.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.

PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0251719.
PR 06-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 03-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-502630/55.
XX N-PSDB; AAK87697.
XX
XX Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognosing disorders of the
PT digestive system, particularly cancer and cancer metastases -
XX
XX Claim 11; SEQ ID NO 1273; 986pp; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a digestive system antigen of
CC the invention.
XX
XX Sequence 90 AA;
SQ
Query Match 81.4%; Score 35; DB 22; Length 90;
Best Local Similarity 71.4%; Pred. No. 14;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Oy 1 HGRVRPH 7
||: |||
Db 73 HGQTRPH 79
RESULT 7
AAU19970
ID AAU19970 standard; Protein; 90 AA.
XX
XX AAU19970;
XX

DT 06-DEC-2001 (first entry)
XX Human liver associated polypeptide #1.
DE
XX Liver associated protein; human; mouse; rabbit; goat; horse; cat; dog;
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW antirheumatic; antiproliferative; cytostatic; cardiac; neuroprotective;
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; tissue regeneration;
anti-infertility.
XX Homo sapiens.
XX WO200155355-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01351.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.

```
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246533.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.

PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0251989.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-457728/49.
XX N-PSDB; AAS31681.
XX
XX Isolated nucleic acid molecule encoding a human liver related protein
XX is used in preventing, treating or ameliorating disorders of the liver
XX particularly cancer of the liver -
XX
XX Claim 11; SEQ ID NO 157; 526pp; English.
XX
XX Sequences AAU1970-AAU20115 represent the liver associated polypeptides
XX of the invention. Liver associated polypeptides and their associated
XX polynucleotides are useful in the diagnosis, treatment and prevention of
XX various types of disorders in e.g. humans, mice, rabbits, goats, horses,
XX cats, dogs, chickens or sheep. A pathological condition can be determined
XX by detecting the presence or absence of a mutation in a liver associated
XX polynucleotide. The treatable disorders include autoimmune diseases such
XX as rheumatoid arthritis, hyperproliferative disorders such as neoplasms
XX of the breast or liver, cardiovascular disorders such as cardiac arrest,
XX cerebrovascular disorders such as cerebral ischaemia, nervous system
XX disorders such as Alzheimer's disease, infections caused by bacteria,
XX viruses and fungi, ocular disorders such as corneal infection, endocrine
XX disorders such as premature labour and infertility, gastrointestinal
XX disorders such as Crohn's disease, renal disorders such as
XX glomerulonephritis and respiratory disorders such as asthma and pleurisy.
XX The polypeptides can also be used to aid wound healing, to prevent skin
XX aging due to sunburn, to maintain organs before transplantation, to
XX regenerate tissues and in chemotaxis.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

Query Match 81.4%; Score 35; DB 22; Length 90;
Best Local Similarity 71.4%; Pred. No. 14;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HGRVRPH 7
Db 73 HGQTRPH 79
||: |||

RESULT 8
ABP40831
ID ABP40831 standard; Protein; 90 AA.
XX
XX AC ABP40831;
XX
XX DT 24-JUL-2002 (first entry)
XX
XX DE Human liver antigen HALSC22, SEQ ID NO:157.
XX
XX Human; liver antigen; liver disorder; hepatic disorder; infection;
XX hepatitis; viral; parasitic; bacterial; fungal; inflammatory condition;
XX cirrhosis; granulomatous hepatitis; toxin damage; drug damage;
```

KW autoimmune disease; Wilson's disease; primary biliary cirrhosis;
KW neoplastic disorder; cancer; tumour; portal hypertension;
KW gastrointestinal disorder; hepatitis; drug screening; gene therapy;
KW chromosome mapping; forensic analysis; antibody preparation;
KW hepatotropic; cytostatic; antiinflammatory; virucide; antibacterial;
KW fungicide; parasiticide; antidote; immunosuppressive.
XX
OS Homo sapiens.
XX
XX US2002042096-A1.
XX
XX 11-APR-2002.
XX
XX 17-JAN-2001; 2001US-0764887.
XX
XX 31-JAN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 28-JUN-2000; 2000US-214886P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220964P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225268P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225757P.
PR 14-AUG-2000; 2000US-225758P.
PR 22-AUG-2000; 2000US-226868P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-229287P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 08-SEP-2000; 2000US-231413P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234997P.
PR 27-SEP-2000; 2000US-235834P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236367P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239935P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241809P.
PR 01-NOV-2000; 2000US-244617P.
PR 17-NOV-2000; 2000US-249299P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.
XX
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX WPI; 2002-381944/41.

DR N-PSDB; ABN90036.
XX
XX New nucleic acid encoding human liver antigens, useful for diagnosis,
PT treatment and prevention of e.g. hepatitis and hepatic cancer, also
PT related polypeptides and antibodies
XX
XX Claim 11; SEQ ID No 157; 181pp; English.
XX
XX The invention relates to 145 novel human liver antigens (ABP40831-
CC ABP40975) and to cDNAs encoding them (ABN90036-ABN90180), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human liver antigen
CC polynucleotides, antibodies against human liver antigens, and the use of
CC liver antigen polynucleotides and polypeptides in diagnosing, treating,
CC prognosing or preventing various disorders of the liver. Such conditions
CC include viral infections (e.g., cytomegalovirus, Epstein-Barr virus,
CC hepatitis A virus, hepatitis B virus and hepatitis C virus), parasitic
CC infections (e.g., Clonorchis sinensis, Echinococcus granulosus and
CC Entamoeba histolytica), and also bacterial and fungal infections. Other
CC disorders that may be treated include inflammatory conditions (e.g.,
CC cirrhosis and granulomatous hepatitis), damage caused by drugs or toxins,
CC autoimmune diseases (e.g. Wilson's disease, primary biliary cirrhosis),
CC neoplastic disorders (e.g., adenomas, haemangiomas and hepatocellular
CC carcinoma), portal hypertension, or gastrointestinal disorders (e.g.,
CC peptic ulcers, gastritis and peritoneal diseases). Liver antigen
CC polypeptides and polynucleotides may also be used in screening for
CC compounds which modulate liver antigen expression or activity. The
CC polynucleotides may further be used for gene therapy, chromosome
CC mapping, in the identification of individuals and in forensic analysis,
CC and the polypeptides may be used as molecular weight markers or to
CC prepare antibodies useful in disease diagnosis, drug targeting and
CC phenotyping. The present sequence represents a human liver antigen of
CC the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO at seqdata.uspto.gov/sequence/
XX
XX Sequence 90 AA;
Query Match 81.4%; Score 35; DB 23; Length 90;
Best Local Similarity 71.4%; Pred. No. 14;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 HGRVRPH 7
||: |||
Db 73 HGQTRPH 79
RESULT 9
AAB63738
ID AAB63738 standard; Protein; 201 AA.
XX
XX AAB63738;
AC
XX
XX 26-MAR-2001 (first entry)
DT
DE
DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1100.
KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine.
XX
OS Homo sapiens.
XX
XX WO200073801-A2.
PN
XX
XX 07-DEC-2000.
PD
XX
XX 26-MAY-2000; 2000WO-US14749.
PF
XX
XX 28-MAY-1999; 99US-0136526.
PR
XX 10-SEP-1999; 99US-0153454.
PR
XX

PA (LUDW-) LUDWIG INST CANCER RES.

PI Obata Y;

DR WPI; 2001-025274/03.

XX Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer -

PS Example 1; Page 693; 799pp; English.

XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP), respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer.

XX SQ Sequence 201 AA;

Query Match 81.4%; Score 35; DB 22; Length 201;

Best Local Similarity 71.4%; Pred. No. 33;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGRVRPH 7

DB 155 HDRIRPH 161

RESULT 10

AAB63739

ID AAB63739 standard; Protein; 210 AA.

XX AC AAB63739;

XX DT 26-MAR-2001 (first entry)

XX Human prostate cancer associated antigen protein sequence SEQ ID NO:1101.

XX Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine.

XX OS Homo sapiens.

XX PN WO200073801-A2.

XX PD 07-DEC-2000.

XX PF 26-MAY-2000; 2000WO-US14749.

XX PR 28-MAY-1999; 99US-0136526.

XX PR 10-SEP-1999; 99US-0153454.

XX PA (LUDW-) LUDWIG INST CANCER RES.

XX PI Obata Y;

XX WPI; 2001-025274/03.

XX Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer -

XX PS Example 1; Page 693-694; 799pp; English.

XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014

CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP), respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer.

XX SQ Sequence 210 AA;

Query Match

Best Local Similarity 81.4%; Score 35; DB 22; Length 210;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGRVRPH 7

DB 152 HDRIRPH 158

RESULT 11

AAB63742

ID AAB63742 standard; Protein; 210 AA.

XX AC AAB63742;

XX DT 26-MAR-2001 (first entry)

XX Human prostate cancer associated antigen protein sequence SEQ ID NO:1104.

XX Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine.

XX OS Homo sapiens.

XX PN WO200073801-A2.

XX PD 07-DEC-2000.

XX PF 26-MAY-2000; 2000WO-US14749.

XX PR 28-MAY-1999; 99US-0136526.

XX PR 10-SEP-1999; 99US-0153454.

XX PA (LUDW-) LUDWIG INST CANCER RES.

XX PI Obata Y;

XX WPI; 2001-025274/03.

XX Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer -

XX Example 1; Page 695; 799pp; English.

XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP), respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer.

XX SQ Sequence 210 AA;

Query Match

Best Local Similarity 81.4%; Score 35; DB 22; Length 210;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
|:|:|:|
Db 154 HDRIRPH 160

RESULT 12
AAU60445
ID AAU60445 standard; Protein; 256 AA.
XX
AC AAU60445;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #21341.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertension; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59609.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID NO 21640; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertension and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 256 AA;
Query Match 81.4%; Score 35; DB 22; Length 256;
Best Local Similarity 85.7%; Pred. No. 43;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
|:|:|:|
Db 209 HGRVRAH 215

RESULT 13
AAB63542
ID AAB63542 standard; Protein; 266 AA.
XX
AC AAB63542;
XX
DT 26-MAR-2001 (first entry)
XX
DE Human gastric cancer associated antigen protein sequence SEQ ID NO:904.
XX
KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine.
XX
OS Homo sapiens.
XX
PN WO200073801-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US14749.
XX
PR 28-MAY-1999; 99US-0136526.
PR 10-SEP-1999; 99US-0153454.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Obata Y;
XX
DR WPI; 2001-025274/03.
XX
PT Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer -
XX
PS Example 1; Page 606-607; 799pp; English.
XX
CC AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP) respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer.
XX
SQ Sequence 266 AA;
Query Match 81.4%; Score 35; DB 22; Length 266;
Best Local Similarity 71.4%; Pred. No. 44;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
|:|:|:|
Db 60 HDRIRPH 66

RESULT 14
ABB92963
ID ABB92963 standard; Protein; 1024 AA.
XX
AC ABB92963;
XX
DT 31-MAY-2002 (first entry)

XX Herbicidally active polypeptide SEQ ID NO 2174.
DE Herbicidal; plant; agriculture; herbicide.
KW Arabidopsis thaliana.
XX WO200210210-A2.
XX 07-FEB-2002.
XX 28-AUG-2001; 2001WO-EP09892.
XX 28-AUG-2001; 2001WO-EP09892.
XX (FARB) BAYER AG.
XX Tietjen K, Weidler M;
XX WPI; 2002-269010/31.
XX Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX Claim 5; SEQ ID NO 2174; 261pp + Sequence Listing; English.
XX The invention relates to identifying target proteins
CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX SQ Sequence 1024 AA;
Query Match 81.4%; Score 35; DB 23; Length 1024;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HGRVRP 6
Db 861 HGRVRP 866
|||||
RESULT 15
ABG01690
ID ABG01690 standard; Protein; 203 AA.
XX AC ABG01690;
XX 13-FEB-2002 (first entry)
XX Novel human diagnostic protein #1681.
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX

PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR N-PSDB; AAS65877.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX Claim 20; SEQ ID NO 32049; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes.
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 203 AA;
Query Match 79.1%; Score 34; DB 22; Length 203;
Best Local Similarity 83.3%; Pred. No. 53;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 2 GRVRPH 7
Db 163 GRVRPH 168
|||||
Search completed: July 16, 2003, 13:12:38
Job time : 25.9024 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:10:13 ; Search time 8.19512 Seconds
(without alignments)
25.132 Million cell updates/sec

Title: US-09-910-582B-3

Perfect score: 43

Sequence: 1 HGRVRPH 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	43	100.0	7	4	US-09-326-718-3
2	34	79.1	605	1	US-08-190-802A-49
3	34	79.1	605	4	US-09-063-950-5
4	34	79.1	605	4	US-08-477-346-49
5	34	79.1	605	4	US-08-473-089-49
6	34	79.1	605	4	US-08-487-072A-49
7	32	74.4	285	4	US-08-992-035A-1
8	31	72.1	97	2	US-09-047-125-25
9	31	72.1	97	3	US-07-736-335E-25
10	31	72.1	139	4	US-09-134-001C-5124
11	30	69.8	30	4	US-08-936-165A-342
12	30	69.8	461	4	US-09-422-936-57
13	30	69.8	803	4	US-09-063-035-2
14	30	69.8	844	4	US-09-422-936-51
15	29	67.4	108	4	US-09-199-637A-417
16	29	67.4	137	3	US-08-444-644-17
17	29	67.4	137	3	US-08-232-246A-17
18	29	67.4	233	3	US-08-444-644-33
19	29	67.4	233	4	US-08-232-246A-33
20	29	67.4	233	3	US-08-444-644-19
21	29	67.4	235	3	US-08-444-644-28
22	29	67.4	235	3	US-08-444-644-42
23	29	67.4	235	4	US-08-232-246A-19
24	29	67.4	235	4	US-08-232-246A-28
25	29	67.4	235	4	US-08-232-246A-42
26	29	67.4	252	4	US-08-858-207A-325
27	29	67.4	321	3	US-09-039-609-4

Sequence 2, Appl
Sequence 7, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 51, Appl
Sequence 51, Appl
Sequence 51, Appl
Sequence 51, Appl
Sequence 1, Appl
Sequence 35, Appl
Sequence 14, Appl
Sequence 27, Appl
Sequence 74, Appl
Sequence 75, Appl
Sequence 74, Appl
Sequence 75, Appl
Sequence 3524, Ap

ALIGNMENTS

RESULT 1

US-09-326-718-3

; Sequence 3, Application US/09326718

; Patent No. 6303573

; GENERAL INFORMATION:

; APPLICANT: Ruoslahti, Erkki

; APPLICANT: Mackenna, Deirdre A.

; TITLE OF INVENTION: Heart Homing Peptides and Methods of

; FILE REFERENCE: P-LJ 3512

; CURRENT APPLICATION NUMBER: US/09/326,718

; CURRENT FILING DATE: 1999-06-07

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 7

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: synthetic construct

US-09-326-718-3

Query Match 100.0%; Score 43; DB 4; Length 7;

Best Local Similarity 100.0%; Pred. No. 1.9e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGRVRPH 7

Db 1 HGRVRPH 7

RESULT 2

US-08-190-802A-49

; Sequence 49, Application US/08190802A

; Patent No. 5519003

; GENERAL INFORMATION:

; APPLICANT: Mochly-Rosen, Daria

; APPLICANT: Ron, Dorit

; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

; NUMBER OF SEQUENCES: 265

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dehlinger & Associates

; STREET: P.O.Box 60850

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306-0850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/190.802A
FILING DATE: 01-FEB-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 8600-0139
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 605 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Insulin-like growth factor binding
protein complex, Fig. 32
US-08-190-802A-49

Query Match 79.1%; Score 34; DB 1; Length 605;
Best Local Similarity 83.3%; Pred. No. 35;
Matches 5; Conservative 1; Mismatches 0; Indels 0;

QY 2 GVRPRH 7
DB 399 GRIRPH 404

RESULT 3
US-09-063-950-5
Sequence 5, Application US/09063950C
Patent No. 6225085
GENERAL INFORMATION:
APPLICANT: Holzman, Douglas A.
TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES
FILE REFERENCE: MEI-019
CURRENT APPLICATION NUMBER: US/09/063,950C
CURRENT FILING DATE: 1998-04-21
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 605
TYPE: PRT
ORGANISM: Papio hamadryas
US-09-063-950-5

Query Match 79.1%; Score 34; DB 4; Length 605;
Best Local Similarity 83.3%; Pred. No. 35;
Matches 5; Conservative 1; Mismatches 0; Indels 0;

QY 2 GVRPRH 7
DB 399 GRIRPH 404

RESULT 4
US-08-477-346-49
Sequence 49, Application US/08477346
Patent No. 6262023
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
US-08-477-346-49

Query Match 79.1%; Score 34; DB 4; Length 605;
Best Local Similarity 83.3%; Pred. No. 35;
Matches 5; Conservative 1; Mismatches 0; Indels 0;

QY 2 GVRPRH 7
DB 399 GRIRPH 404

NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,346
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,072
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 605 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Insulin-like growth factor binding
protein complex, Fig. 32
US-08-477-346-49

Query Match 79.1%; Score 34; DB 4; Length 605;
Best Local Similarity 83.3%; Pred. No. 35;
Matches 5; Conservative 1; Mismatches 0; Indels 0;

QY 2 GVRPRH 7
DB 399 GRIRPH 404

RESULT 5
US-08-473-089-49
Sequence 49, Application US/08473089
Patent No. 6342368
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/473,089

;
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 2550-0025.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 605 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Insulin-like growth factor binding
; INDIVIDUAL ISOLATE: protein complex, Fig. 32
US-08-473-089-49

Query Match 79.1%; Score 34; DB 4; Length 605;
Best Local Similarity 83.3%; Pred. No. 35;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVRPH 7
II:III
DB 399 GRIRPH 404

RESULT 6
US-08-487-072A-49
; Sequence 49, Application US/08487072A
; Patent No. 6423684
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Theof
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,072A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 2550-0025.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 605 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:

;
; INDIVIDUAL ISOLATE: Insulin-like growth factor binding
; INDIVIDUAL ISOLATE: protein complex, Fig. 32
US-08-487-072A-49

Query Match 79.1%; Score 34; DB 4; Length 605;
Best Local Similarity 83.3%; Pred. No. 35;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVRPH 7
II:III
DB 399 GRIRPH 404

RESULT 7
US-08-992-035A-1
; Sequence 1, Application US/08992035A
; Patent No. 6242179
; GENERAL INFORMATION:
; APPLICANT: Shah, Purvi
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: HUMAN PHOSPHATASES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/992,035A
; FILING DATE: December 17, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0433 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 285 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BLADNOT06
; CLONE: 1719418
US-08-992-035A-1

Query Match 74.4%; Score 32; DB 4; Length 285;
Best Local Similarity 83.3%; Pred. No. 39;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVRPH 7
II:III
DB 125 GRLRPH 130

RESULT 8
US-09-047-125-25
; Sequence 25, Application US/09047125
; Patent No. 5976787

QY 1 HGRVRPH 7
||:|:
Db 86 HGKVRPY 92

RESULT 11

US-08-936-165A-342
; Sequence 342, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,165A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,032
; FILING DATE: 24-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 342:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-936-165A-342

Query Match 69.8%; Score 30; DB 4; Length 30;
Best Local Similarity 83.3%; Pred. No. 8.3;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GRVRPH 7
||:|:
Db 24 GRVNP 29

RESULT 12

US-09-422-936-57
; Sequence 57, Application US/09422936
; Patent No. 6465213
; GENERAL INFORMATION:

; APPLICANT: Ekstrand, Jonas
; TITLE OF INVENTION: NEW NUCLEOTIDE SEQUENCES
; FILE REFERENCE: 06275-165002
; CURRENT APPLICATION NUMBER: US/09/422,936
; CURRENT FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: US 09/242,608
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: PCT/SE98/01947
; PRIOR FILING DATE: 1998-10-27
; PRIOR APPLICATION NUMBER: SWEDEN 9703914-2
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: SWEDEN 9800864-2
; PRIOR FILING DATE: 1998-03-16
; PRIOR APPLICATION NUMBER: SWEDEN 9802575-2
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-422-936-57

Query Match 69.8%; Score 30; DB 4; Length 461;
Best Local Similarity 71.4%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
| | | | |
Db 39 HSRVPPH 45

RESULT 13

US-09-063-035-2
; Sequence 2, Application US/09063035
; Patent No. 6160091
; GENERAL INFORMATION:
; APPLICANT: PEUKERT, Karen; HAENEL, Frank; and EILERS,
; APPLICANT: Martin
; TITLE OF INVENTION: Myc-binding zinc finger proteins,
; their preparation and their use
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Keil & Weinkauff
; STREET: 1101 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.2 Mb storage
; COMPUTER: IBM AT-compatible, 80486 processor
; OPERATING SYSTEM: MS-DOS version 6.1
; SOFTWARE: WordPerfect version 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/063,035
; FILING DATE: 21-APR-1998
; CLASSIFICATION: 514
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 803 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-063-035-2

Query Match 69.8%; Score 30; DB 4; Length 803;
Best Local Similarity 57.1%; Pred. No. 3.1e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
| | | | |
Db 580 HDNIRPH 586

Search completed: July 16, 2003, 13:17:06
Job time : 9.19512 secs

RESULT 14

US-09-422-936-51
; Sequence 51, Application US/09422936
; Patent No. 6465213
; GENERAL INFORMATION:
; APPLICANT: Ekstrand, Jonas
; TITLE OF INVENTION: NEW NUCLEOTIDE SEQUENCES
; FILE REFERENCE: 06275-165002
; CURRENT APPLICATION NUMBER: US/09/422,936
; CURRENT FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: US 09/242,608
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: PCT/SE98/01947
; PRIOR FILING DATE: 1998-10-27
; PRIOR APPLICATION NUMBER: SWEDEN 9703914-2
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: SWEDEN 9800864-2
; PRIOR FILING DATE: 1998-03-16
; PRIOR APPLICATION NUMBER: SWEDEN 9802575-2
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 844
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-422-936-51

Query Match 69.8%; Score 30; DB 4; Length 844;
Best Local Similarity 71.4%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
| | | | |
Db 39 HSRVPPH 45

RESULT 15

US-09-199-637A-417
; Sequence 417, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 417
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-417

Query Match 67.4%; Score 29; DB 4; Length 108;
Best Local Similarity 71.4%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
| | | | |
Db 28 HPRPRPH 34

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:15:24 ; Search time 13.3171 Seconds
(without alignments)
62.425 Million cell updates/sec

Title: US-09-910-582B-3
Perfect score: 43
Sequence: 1 HGRVRPH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues
Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published_Applications_AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	100.0	7	9	US-09-910-582B-3
2	43	100.0	7	10	US-09-782-650-2
3	35	81.4	90	9	US-10-073-961-157
4	35	81.4	90	10	US-09-764-887-157
5	34	79.1	605	10	US-09-782-980-62
6	32	74.4	86	10	US-09-867-550-486
7	32	74.4	308	9	US-09-934-455-30
8	32	74.4	403	9	US-09-738-626-5159
9	32	74.4	507	9	US-10-156-761-11305
10	32	74.4	701	9	US-10-156-761-8059
11	32	74.4	1171	9	US-10-137-666A-96
12	32	74.4	1198	9	US-10-197-666A-98
13	32	74.4	1198	12	US-10-001-215-2
14	32	74.4	1736	10	US-09-919-497-98
15	31	72.1	108	9	US-09-764-891-4485
16	31	72.1	414	10	US-09-820-893-69
17	31	72.1	457	10	US-09-888-615-110
18	31	72.1	480	10	US-09-820-893-108
19	31	72.1	536	10	US-09-864-761-36148

20	31	72.1	793	9	US-10-156-761-9605	Sequence 9605, Ap
21	30.5	70.9	220	9	US-10-189-346-93	Sequence 93, Appl
22	30.5	70.9	246	9	US-10-189-346-70	Sequence 70, Appl
23	30.5	70.9	246	9	US-10-189-346-90	Sequence 90, Appl
24	30	69.8	18	9	US-10-300-616-37	Sequence 37, Appl
25	30	69.8	30	10	US-09-939-980-342	Sequence 342, App
26	30	69.8	44	10	US-09-864-761-46951	Sequence 46951, A
27	30	69.8	44	10	US-09-864-761-48177	Sequence 48177, A
28	30	69.8	70	10	US-09-864-761-44516	Sequence 44516, A
29	30	69.8	124	9	US-10-106-698-6925	Sequence 6925, Ap
30	30	69.8	143	10	US-09-739-438-4	Sequence 4, Appl
31	30	69.8	176	9	US-10-156-761-10802	Sequence 10802, A
32	30	69.8	190	9	US-10-156-761-11426	Sequence 11426, A
33	30	69.8	240	9	US-10-156-761-11975	Sequence 11975, A
34	30	69.8	286	9	US-10-156-761-14671	Sequence 14671, A
35	30	69.8	379	9	US-09-738-626-4255	Sequence 4255, Ap
36	30	69.8	513	10	US-09-833-745-59	Sequence 59, Appl
37	30	69.8	513	10	US-09-833-745-60	Sequence 60, Appl
38	30	69.8	513	10	US-09-833-745-61	Sequence 61, Appl
39	30	69.8	583	10	US-09-810-264-38	Sequence 38, Appl
40	30	69.8	664	9	US-10-156-761-13352	Sequence 13352, A
41	30	69.8	749	9	US-10-099-352-40	Sequence 40, Appl
42	30	69.8	844	10	US-09-875-724-8	Sequence 8, Appl
43	30	69.8	1060	10	US-09-801-368-410	Sequence 410, App
44	30	69.8	3472	9	US-10-027-806-4	Sequence 4, Appl
45	30	69.8	3472	9	US-10-034-623-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-910-582B-3
; Sequence 3, Application US/09910582B
; Publication No. US20030045476A1
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; TITLE OF INVENTION: Heart Homing Conjugates
; FILE REFERENCE: P-LJ 4857
; CURRENT APPLICATION NUMBER: US/09/910,582B
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 09/326,718
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-09-910-582B-3

Query Match 100.0%; Score 43; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGRVRPH 7
| | | | |
Db 1 HGRVRPH 7

RESULT 2
US-09-782-650-2
; Sequence 2, Application US/09782650
; Patent No. US20020019350A1
; GENERAL INFORMATION:
; APPLICANT: Levine, Arnold J.
; APPLICANT: Mitterer, Artur
; APPLICANT: Falkner, Falko-Gunter
; APPLICANT: Scheiflinger, Friedrich
; APPLICANT: Dörner, Friedrich

us-09-910-582b-3.rapb

Wed Jul 16 13:41:35 2003

APPLICANT: Edwards Lifesciences Corporation
TITLE OF INVENTION: Targeted Angiogenesis
FILE REFERENCE: 20553D-000611US
CURRENT APPLICATION NUMBER: US/09/782,650
CURRENT FILING DATE: 2001-02-12
PRIOR APPLICATION NUMBER: US 09/324,079
PRIOR FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: US 09/327,045
PRIOR FILING DATE: 1999-06-07
PRIOR APPLICATION NUMBER: PCT/US00/14988
PRIOR FILING DATE: 2000-05-31
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:targeting
OTHER INFORMATION: molecule.
US-09-782-650-2

Query Match 100.0%; Score 43; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
Db 1 HGRVRPH 7

RESULT 3
US-10-073-961-157
Sequence 157, Application US/10073961
Publication No. US20030077602A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA113C1
CURRENT APPLICATION NUMBER: US/10/073,961
CURRENT FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: 09/764,887
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/225,757
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226,868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216,647
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,267
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/216,880
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,270
PRIOR FILING DATE: 2000-08-14

PRIOR APPLICATION NUMBER: 60/251,869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235,834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234,274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234,223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228,924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224,518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249,299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239,937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241,787

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14

; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08

Query Match 81.4%; Score 35; DB 9; Length 90;
Best Local Similarity 71.4%; Pred. No. 14;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
Db 73 HGQTRPH 79
11: 111
73 HGQTRPH 79

RESULT 4
US-09-764-887-157
; Sequence 157, Application US/09764887
; Patent No. US20020042096A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P4113
; CURRENT APPLICATION NUMBER: US/09/764,887
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 157
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (16)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-887-157

Query Match 81.4%; Score 35; DB 10; Length 90;
Best Local Similarity 71.4%; Pred. No. 14;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
Db 73 HGQTRPH 79
11: 111
73 HGQTRPH 79

RESULT 5
US-09-782-980-62
; Sequence 62, Application US/09782980
; Patent No. US20020072089A1
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran M.
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Busfield, Samantha J.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Gu, Wei
; APPLICANT: White, David
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL ITALY, LOR-2, STRIFE, TRASH, BDSF, LRSG, AND
; TITLE OF INVENTION: STMSST PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-121CP

OTHER INFORMATION: wherein Xaa may be any one of Arg or Ile or Lys or Thr

US-09-867-550-486

Query Match 74.4%; Score 32; DB 10; Length 86;
Best Local Similarity 71.4%; Pred. No. 48;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
DB 34 HGRMRDH 40

RESULT 7

US-09-934-455-30

Sequence 30, Application US/09934455
Publication No. US20030121070A1

GENERAL INFORMATION:

APPLICANT: Adam, Luc
APPLICANT: Creelman, Robert
APPLICANT: Dubell, Arnold
APPLICANT: Heard, Jacqueline
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Keddie, James
APPLICANT: Pilgrim, Marsha
APPLICANT: Ratcliffe, Oliver
APPLICANT: Reuber, Lynne
APPLICANT: Riechmann, Jose Luis
APPLICANT: Yu, Guo-Liang
APPLICANT: Pineda, Omaira

TITLE OF INVENTION: Genes for Modifying Plant Traits IV

FILE REFERENCE: MBI-0025

CURRENT APPLICATION NUMBER: US/09/934,455

CURRENT FILING DATE: 2001-08-22

PRIOR APPLICATION NUMBER: 60/227439

PRIOR FILING DATE: 2000-08-22

PRIOR APPLICATION NUMBER: MBI-0022

PRIOR FILING DATE: 2001-11-16

PRIOR APPLICATION NUMBER: MBI-0023

PRIOR FILING DATE: 2001-04-17

NUMBER OF SEQ ID NOS: 516

SOFTWARE: Patent in version 3.1

SEQ ID NO 30

LENGTH: 308

TYPE: PRT

ORGANISM: Arabidopsis thaliana

US-09-934-455-30

Query Match 74.4%; Score 32; DB 9; Length 308;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
DB 127 HGNVTPH 133

RESULT 8

US-09-738-626-5159

Sequence 5159, Application US/09738626
Publication No. US20020197605A1

GENERAL INFORMATION:

APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-125

CURRENT APPLICATION NUMBER: US/09/782,980

CURRENT FILING DATE: 2001-02-13

PRIOR APPLICATION NUMBER: PCT/US00/021125

PRIOR FILING DATE: 2000-01-27

PRIOR APPLICATION NUMBER: 09/448,076

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: 09/276,400

PRIOR FILING DATE: 1999-03-25

PRIOR APPLICATION NUMBER: 60/117,580

PRIOR FILING DATE: 1999-01-27

PRIOR APPLICATION NUMBER: 09/014,195

PRIOR FILING DATE: 1998-01-27

PRIOR APPLICATION NUMBER: 09/014,348

PRIOR FILING DATE: 1998-01-27

PRIOR APPLICATION NUMBER: 09/086,892

PRIOR FILING DATE: 1998-05-29

PRIOR APPLICATION NUMBER: 09/296,208

PRIOR FILING DATE: 1999-04-21

PRIOR APPLICATION NUMBER: 09/063,950

PRIOR FILING DATE: 1998-04-21

PRIOR APPLICATION NUMBER: 09/561,381

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 09/561,810

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 09/087,121

PRIOR FILING DATE: 1998-05-29

PRIOR APPLICATION NUMBER: 09/672,721

PRIOR FILING DATE: 2000-09-28

PRIOR APPLICATION NUMBER: 09/049,799

PRIOR FILING DATE: 1998-03-27

NUMBER OF SEQ ID NOS: 176

SOFTWARE: Patent in Ver. 2.0

SEQ ID NO 62

LENGTH: 605

TYPE: PRT

ORGANISM: Papio hamadryas

US-09-782-980-62

Query Match 79.1%; Score 34; DB 10; Length 605;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVRPH 7
DB 399 GRIRPH 404

RESULT 6

US-09-867-550-486

Sequence 486, Application US/09867550
Patent No. US20020082206A1

GENERAL INFORMATION:

APPLICANT: Leach, Martin D.
APPLICANT: Mehraban, Fuad,
APPLICANT: Conley, Pamela
APPLICANT: Law, Debbie
APPLICANT: Topper, James

TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and

FILE REFERENCE: 21402-013 (Cura-313)

CURRENT APPLICATION NUMBER: US/09/867,550

CURRENT FILING DATE: 2001-09-20

PRIOR APPLICATION NUMBER: USSN 60/208,427

PRIOR FILING DATE: 2000-05-30

NUMBER OF SEQ ID NOS: 2125

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 486

LENGTH: 86

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: VARIANT

LOCATION: (13)

; CURRENT APPLICATION NUMBER: US/09/738,626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 5159
 ; LENGTH: 403
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-5159

Query Match 74.4%; Score 32; DB 9; Length 403;
 Best Local Similarity 83.3%; Pred. No. 2.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HGRVRP 6
 Db 208 HGRVKP 213
 ||||:|

RESULT 9
 US-10-156-761-11305
 ; Sequence 11305, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 11305
 ; LENGTH: 507
 ; TYPE: PRT
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-11305

Query Match 74.4%; Score 32; DB 9; Length 507;
 Best Local Similarity 83.3%; Pred. No. 3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HGRVRP 6
 Db 55 HGRLRP 60
 ||||:|

RESULT 10
 US-10-156-761-8059
 ; Sequence 8059, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRO

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 8059
 ; LENGTH: 701
 ; TYPE: PRT
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-8059

Query Match 74.4%; Score 32; DB 9; Length 701;
 Best Local Similarity 83.3%; Pred. No. 4.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HGRVRP 6
 Db 419 HGRMRP 424
 ||||:|

RESULT 11
 US-10-197-666A-96
 ; Sequence 96, Application US/10197666A
 ; Publication No. US20030092037A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ASAHI KASEI KABUSIKI KAISYA
 ; TITLE OF INVENTION: Elkl phosphorylation related gene
 ; FILE REFERENCE: PH-1548US
 ; CURRENT APPLICATION NUMBER: US/10/197,666A
 ; CURRENT FILING DATE: 2002-11-18
 ; PRIOR APPLICATION NUMBER: JP 2001-218204
 ; PRIOR FILING DATE: 2001-07-18
 ; PRIOR APPLICATION NUMBER: JP 2001-263450
 ; PRIOR FILING DATE: 2001-08-31
 ; PRIOR APPLICATION NUMBER: JP 2002-012176
 ; PRIOR FILING DATE: 2002-01-21
 ; PRIOR APPLICATION NUMBER: US 60/305,884
 ; PRIOR FILING DATE: 2001-07-18
 ; PRIOR APPLICATION NUMBER: US 60/316,304
 ; PRIOR FILING DATE: 2001-09-04
 ; PRIOR APPLICATION NUMBER: US 60/350,027
 ; PRIOR FILING DATE: 2002-01-23
 ; NUMBER OF SEQ ID NOS: 156
 ; SOFTWARE: PatentIn ver. 2.1
 ; SEQ ID NO 96
 ; LENGTH: 1171
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-197-666A-96

Query Match 74.4%; Score 32; DB 9; Length 1171;
 Best Local Similarity 71.4%; Pred. No. 7e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 HGRVRPH 7
 Db 1053 HGTVPH 1059
 ||||:|

RESULT 12
 US-10-197-666A-98
 ; Sequence 98, Application US/10197666A
 ; Publication No. US20030092037A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ASAHI KASEI KABUSIKI KAISYA
 ; TITLE OF INVENTION: Elkl phosphorylation related gene
 ; FILE REFERENCE: PH-1548US
 ; CURRENT APPLICATION NUMBER: US/10/197,666A
 ; CURRENT FILING DATE: 2002-11-18

; PRIOR APPLICATION NUMBER: JP 2001-218204
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: JP 2001-263450
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: JP 2002-012176
; PRIOR FILING DATE: 2002-01-21
; PRIOR APPLICATION NUMBER: US 60/305,884
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/316,304
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 60/350,027
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 1198
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-197-666A-98

Query Match 74.4%; Score 32; DB 9; Length 1198;
Best Local Similarity 71.4%; Pred. No. 7.2e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
||| ||
DB 1080 HGTVHPH 1086

RESULT 13
US-10-001-215-2
; Sequence 2, Application US/10001215
; Patent No. US20020147323A1
; GENERAL INFORMATION:
; APPLICANT: Bandaru Rajasekhara
; APPLICANT: Rosana Kapeller-Libermann
; TITLE OF INVENTION: 16224 and 69611, NOVEL HUMAN KINASES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: MNI-209
; CURRENT APPLICATION NUMBER: US/10/001,215
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/250,917
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1198
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-001-215-2

Query Match 74.4%; Score 32; DB 12; Length 1198;
Best Local Similarity 71.4%; Pred. No. 7.2e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
||| ||
DB 1080 HGTVHPH 1086

RESULT 14
US-09-919-497-98
; Sequence 98, Application US/09919497
; Patent No. US20020106662A1
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7725
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 98
; LENGTH: 1736
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-919-497-98

Query Match 74.4%; Score 32; DB 10; Length 1736;
Best Local Similarity 83.3%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGRVRP 6
||| ||
DB 1156 HGRURP 1161

RESULT 15

US-09-764-891-4485
; Sequence 4485, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4485
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (49)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (51)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (73)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (93)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-4485

Query Match 72.1%; Score 31; DB 9; Length 108;
Best Local Similarity 83.3%; Pred. No. 94;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGRVRP 6
||| ||
DB 67 HGRTRP 72

Search completed: July 16, 2003, 13:35:56
Job time : 14.3171 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:09:18 ; Search time 9.21951 Seconds
(without alignments)
72.991 Million cell updates/sec

Title: US-09-910-582B-3

Perfect score: 43

Sequence: 1 HGRVRPH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 9613422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	100.0	118	2 T07905	low-carbon dioxide
2	37	86.0	162	2 D82319	probable sprt prot
3	35	81.4	229	2 F75482	transcription regu
4	35	81.4	442	2 T34714	hypothetical prote
5	35	81.4	1024	2 G71434	probable limonene
6	34	79.1	68	2 S33581	hypothetical prote
7	34	79.1	242	2 T15361	hypothetical prote
8	34	79.1	250	2 T35927	conserved hypothet
9	34	79.1	355	2 G83254	probable secretion
10	34	79.1	475	2 F64151	hypothetical prote
11	34	79.1	497	2 S22708	homeotic protein e
12	34	79.1	605	2 JC5239	insulin-like growt
13	34	79.1	605	2 A41915	insulin-like growt
14	33	76.7	206	2 C83590	probable transcrip
15	33	76.7	711	2 S68443	double-stranded RN
16	32.5	75.6	352	2 A95931	probable threonine
17	32	74.4	123	2 H72709	hypothetical prote
18	32	74.4	156	1 B64021	hypothetical prote
19	32	74.4	249	2 A83963	hypothetical prote
20	32	74.4	318	2 JE0199	phosphatidic acid
21	32	74.4	318	2 H89632	protein Fl36.5 [i
22	32	74.4	341	2 T16951	hypothetical prote
23	32	74.4	377	2 AE0715	probable membrane
24	32	74.4	409	2 JC7293	activated T cell t
25	32	74.4	512	2 T38422	probable chromatin
26	32	74.4	654	2 PC7085	nuclear factor of
27	32	74.4	1189	2 T17088	homeodomain-inte
28	32	74.4	1325	2 S16129	dynein-associated
29	32	74.4	1736	2 A47747	tight junction pro

30	31	72.1	64	2 S53450	annexins - sea urc
31	31	72.1	74	2 F97551	hypothetical prote
32	31	72.1	170	2 S32578	ribosomal protein
33	31	72.1	171	2 T01410	ribosomal protein
34	31	72.1	172	2 S32579	ribosomal protein
35	31	72.1	175	2 T02163	ribosomal protein
36	31	72.1	183	2 S34122	ribosomal protein
37	31	72.1	184	1 R5HU22	ribosomal protein
38	31	72.1	184	1 R5RT17	ribosomal protein
39	31	72.1	224	2 F70798	hypothetical prote
40	31	72.1	282	2 C72660	hypothetical prote
41	31	72.1	315	2 A91118	probable ferrichro
42	31	72.1	315	2 H85962	probable ferrichro
43	31	72.1	325	2 D75279	conserved hypothet
44	31	72.1	382	2 AH2056	hypothetical prote
45	31	72.1	388	2 E72207	GTP cyclohydrolase

ALIGNMENTS

5/14/99

RESULT 1

T07905

Low-carbon dioxide inducible protein 3 - Chlamydomonas reinhardtii

C:Species: Chlamydomonas reinhardtii

C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 08-Oct-1999.

C:Accession: T07905

R(Somanchi) A.; Handley, E.R.; Moroney, J.V.

submitted to the EMBL Data Library, July 1997

A:Description: Chlamydomonas reinhardtii cDNAs upregulated in low-carbon dioxide cond

A:Reference number: Z16205

A:Accession: T07905

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-118 <SOM>

A:Cross-references: EMBL:AF015661; NID:g2367584; PIDN:AAB69691.1; PID:g2367585

C:Genetics:

A:Gene: lci3

Query Match 100.0%; Score 43; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1 HGRVRPH 7
Db	34 HGRVRPH 40

RESULT 2

D82319

probable sprt protein VC0471 [imported] - Vibrio cholerae (strain N16961 serogroup O1

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: D82319

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers

l. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406: 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: D82319

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-162 <HEI>

A:Cross-references: GB:AE004133; GB:AE003852; NID:g9654889; PIDN:AAF93644.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC0471

A:Map position: 1

C:Superfamily: hypothetical protein H1173

Query Match 86.0%; Score 37; DB 2; Length 162;
Best Local Similarity 85.7%; Pred. No. 3.6;

Wed Jul 16 13:41:36 2003

```

R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Ferry, N.; G
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomen
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.;
C.; Chalwatzis, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis t
A:Reference number: A71400; MUID:98121113; PMID:9461215
A:Accession: G71434
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1024 <BEV>
A:Cross-references: GB:297341; NID:g2244991; PID:e327011; PID:g2245029
C:Genetics:
A:Map position: 4COP9-4G3845

Query Match      81.4%; Score 35; DB 2; Length 1024;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HGRVRP 6
      |||||
Db      861 HGRVRP 866

RESULT 6
S33581
hypothetical protein 3 - Azorhizobium caulinodans
C:Species: Azorhizobium caulinodans
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 11-May-2000
C:Accession: S33581
R:Stigter, J.; Schneider, M.; de Bruijn, F.J.
Mol. Plant Microbe Interact. 6, 238-252, 1993
A:Title: Azorhizobium caulinodans nitrogen fixation (nif/fix) gene regulation: mutagenesis
A:Reference number: S33579; MUID:93229807; PMID:8471796
A:Accession: S33581
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-68 <STI>
A:Cross-references: EMBL:X69959; NID:g311387; PIDN:CAA49583.1; PID:g311390

Query Match      79.1%; Score 34; DB 2; Length 68;
Best Local Similarity 85.7%; Pred. No. 6.2;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 HGRVRP 7
      |||||
Db      38 HGRVRP 44

RESULT 7
T15361
hypothetical protein B0563.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15361
R:Favella, T.
Submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid B0563.
A:Reference number: Z18336
A:Accession: T15361
A:Molecule type: DNA
A:Residues: 1-242 <FAV>
A:Cross-references: EMBL:U28740; NID:g861281; PID:g861287; PIDN:AAA68322.1; CESP:B056
C:Genetics:
A:Experimental source: Strain Bristol N2
A:Gene: CESP:B0563.2
A:Introns: 34/3; 76/3; 118/3; 143/3; 187/3
Query Match      79.1%; Score 34; DB 2; Length 242;
Best Local Similarity 71.4%; Pred. No. 22;

QY      1 HGRVRP 6
      |||||
Db      349 HGRVRP 354

RESULT 5
G71434
probable limonene cyclase - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
C:Accession: G71434

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: F75482
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-229 <WHI>
A:Cross-references: GB:AE001929; GB:AE000513; NID:g6458437; PIDN:AAF10313.1; PID:g645844
A:Experimental source: strain R1.
C:Genetics:
A:Gene: DR0735
A:Map position: 1

Query Match      81.4%; Score 35; DB 2; Length 229;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 GRVRP 7
      |||||
Db      148 GRVRP 153

RESULT 4
T34714
hypothetical protein SC1C3.22 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 04-Mar-2000
C:Accession: T34714
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
Submitted to the EMBL Data Library, May 1998
A:Reference number: Z21554
A:Accession: T34714
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-442 <OLI>
A:Cross-references: EMBL:AL023702; PIDN:CAA19246.1; GSPDB:GN00070; SCOEDB:SC1C3.22
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC1C3.22
C:Superfamily: Streptomyces coelicolor hypothetical protein SC1C3.22

Query Match      81.4%; Score 35; DB 2; Length 442;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HGRVRP 6
      |||||
Db      349 HGRVRP 354

RESULT 5
G71434
probable limonene cyclase - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
C:Accession: G71434
```

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HGRVRPH 7
 |||||:
 Db 225 HGRLRPY 231

RESULT 8

T35927
 conserved hypothetical protein SC9B5.07 - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T35927
 R:Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, January 1999
 A:Reference number: Z21551
 A:Accession: T35927
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-250 <SEE>
 A:Cross-references: EMBL:AL035206; PIDN:CAA22749.1; GSPDB:GN00070; SCOEDB:SC9B5.07
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SC9B5.07

Query Match 79.1%; Score 34; DB 2; Length 250;
 Best Local Similarity 83.3%; Pred. No. 22;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVRVRPH 7
 |||||:
 Db 222 GRIRPH 227

RESULT 9

G83254
 probable secretion protein PA3136 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: G83254
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: G83254
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-355 <STO>
 A:Cross-references: GB:AE004737; GB:AE004091; NID:99949239; PIDN:AAG06524.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA3136
 C:Superfamily: multidrug resistance protein A; lipoyl/biotin-binding homology

Query Match 79.1%; Score 34; DB 2; Length 355;
 Best Local Similarity 83.3%; Pred. No. 32;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGRVRPH 6
 |||||:
 Db 337 HGRIRP 342

RESULT 10

F64151
 hypothetical protein HI0409 - Haemophilus influenzae (strain Rd KW20)
 C:Species: Haemophilus influenzae
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 29-Sep-1999
 C:Accession: F64151
 R:Flieschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J

, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
 Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630; PMID:7542800
 A:Accession: F64151
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-475 <TIGR>
 A:Cross-references: GB:U32724; GB:I42023; NID:g1573378; PIDN:AAC22068.1; PID:g1573382
 A:Note: best homolog was a hypothetical protein from Escherichia coli
 C:Superfamily: hypothetical protein HI0409

Query Match 79.1%; Score 34; DB 2; Length 475;
 Best Local Similarity 83.3%; Pred. No. 42;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVRVRPH 7
 |||||:
 Db 343 GRIRPH 348

RESULT 11

S22708
 homeotic protein ems - fruit fly (Drosophila melanogaster)
 N:Alternate names: empty spiracles protein
 C:Species: Drosophila melanogaster
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Oct-1999
 C:Accession: S22708; A34218
 R:Waldorf, U.; Gehring, W.J.
 EMBO J. 11, 2247-2259, 1992
 A:Title: Empty spiracles, a gap gene containing a homeobox involved in Drosophila hea
 A:Reference number: S22708; MUID:92289690; PMID:1376248
 A:Accession: S22708
 A:Molecule type: DNA
 A:Residues: 1-497 <WAL>
 A:Cross-references: EMBL:X66270; NID:g8819; PIDN:CAA46985.1; PID:g8820
 R:Dalton, D.; Chadwick, R.; McGinnis, W.
 Genes Dev. 3, 1940-1956, 1989
 A:Title: Expression and embryonic function of empty spiracles: a Drosophila homeo box
 A:Reference number: A34218; MUID:90152339; PMID:2576012
 A:Accession: A34218
 A:Molecule type: mRNA
 A:Residues: 1-3, 'T', '5'-101,105-130, 'Q', '132-157, 'TR', '160-303, 'DV', '306-497 <DAL>
 A:Cross-references: GB:X51653; NID:g7944; PIDN:CAA35965.1; PID:g671861
 C:Genetics:
 A:Gene: FlyBase:ems
 A:Cross-references: FlyBase:FBgn0000576
 A:Introns: 382/1
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:392-448/Domain: homeobox homology <HGX>

Query Match 79.1%; Score 34; DB 2; Length 497;
 Best Local Similarity 71.4%; Pred. No. 44;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
 |||||:
 Db 371 HGRIRPH 377

RESULT 12

JC5239
 insulin-like growth factor acid-labile chain - baboon
 C:Species: Papio sp. (baboon)
 C:Date: 17-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-May-1997
 C:Accession: JC5239
 R:Delhanty, P.; Baxter, R.C.
 Biochem. Biophys. Res. Commun. 227, 897-902, 1996
 A:Title: The cloning and expression of the baboon acid-labile subunit of the insulin-
 A:Reference number: JC5239; MUID:97040714; PMID:8886027
 A:Contents: liver

A:Accession: JC5239

A:Molecule type: mRNA

A:Residues: 1-605

C:Comment: This factor is structurally related to proinsulin and have insuline-like meta

Query Match 79.1%; Score 34; DB 2; Length 605;

Best Local Similarity 83.3%; Pred. No. 54;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRVRPH 7

||:||||

Db 399 GRIRPH 404

RESULT 13

A41915

Insulin-like growth factor-binding complex acid-labile chain precursor - human

N:Alternate names: Acid-labile Subunit (ALS)

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 05-Nov-1999

C:Accession: A41915

R:Leong, S.R.; Baxter, R.C.; Camerato, T.; Dai, J.; Wood, W.I.

Mol. Endocrinol. 6, 870-876, 1992

A:Title: Structure and functional expression of the acid-labile subunit of the insulin-l

A:Reference number: A41915; MUID:92357025; PMID:1379671

A:Accession: A41915

A>Status: preliminary

A:Molecule type: mRNA; protein

A:Residues: 1-605 <LEO>

A:Cross-references: GB:M86826; NID:gl84807; PIDN:AAA36047.1; PID:gi848408

A:Experimental source: liver

A:Note: sequence extracted from NCBI backbone (NCBIP:110171)

C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology

F:75-98/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F:99-122/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:123-146/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:147-170/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F:171-194/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F:195-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F:219-242/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

F:243-266/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>

F:267-290/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>

F:291-314/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>

F:315-338/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>

F:339-362/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>

F:363-386/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>

F:387-410/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>

F:411-434/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>

F:435-458/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR16>

F:459-482/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR17>

F:483-506/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR18>

F:507-529/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR19>

Query Match 79.1%; Score 34; DB 2; Length 605;

Best Local Similarity 83.3%; Pred. No. 54;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GRVRPH 7

||:||||

Db 399 GRIRPH 404

RESULT 14

C83590

Probable transcription regulator PA0436 [imported] - Pseudomonas aeruginosa (strain PA01

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: C83590

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: C83590

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-206 <STO>

A:Cross-references: GB:AE004481; GB:AE004091; NID:g9946293; PIDN:AAG03825.1; GSPDB:GN

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA0436

Query Match 76.7%; Score 33; DB 2; Length 206;

Best Local Similarity 71.4%; Pred. No. 29;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 HGRVRPH 7

||:||||

Db 97 HSTRPH 103

RESULT 15

S68443

N:Alternate names: double-stranded RNA-specific editase

C:Species: Rattus norvegicus (Norway rat)

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Jun-1999

C:Accession: S68443; S78526

R:Melcher, T.; Maas, S.; Herb, A.; Sprengel, R.; Seeburg, P.H.; Higuchi, M.

Nature 379, 460-464, 1996

A:Title: A mammalian RNA editing enzyme.

A:Reference number: S68443; MUID:96158881; PMID:8559253

A:Accession: S68443

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-711 <MEL>

A:Cross-references: EMBL:U03534

R:Seeburg, P.H.

submitted to the EMBL Data Library, December 1995

A:Reference number: S78526

A:Accession: S78526

A:Molecule type: mRNA

A:Residues: 1-454, 'R', 456-711 <SEP>

A:Cross-references: EMBL:U03534; NID:gl177852; PIDN:AAA96755.1; PID:gl177853

C:Genetics:

A:Gene: REDI

C:Superfamily: double-stranded RNA-specific adenosine deaminase; double-stranded RNA-

C:Keywords: hydrolase; metalloprotein; RNA binding; RNA editing; zinc

F:77-145/Domain: double-stranded RNA-binding repeat homology <DSR1>

F:248-299/Domain: double-stranded RNA-binding repeat homology <DSR3>

F:394,451,526/Binding site: zinc (His, Cys, Cys) #status predicted

Query Match 76.7%; Score 33; DB 2; Length 711;

Best Local Similarity 71.4%; Pred. No. 1e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HGRVRPH 7

||:||||

Db 650 HGRVPPH 656

Search completed: July 16, 2003, 13:16:12

Job time : 11.2195 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:04:48 ; Search time 4.78049 Seconds
(without alignments)
60.733 Million cell updates/sec

Title: US-09-910-582B-3
Perfect score: 43
Sequence: 1 HGRVRPH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	35	81.4	731	1 NGP1_HUMAN	Q13823 homo sapien
2	34	79.1	68	1 YRP3_AZOCA	P33988 azorhizobiu
3	34	79.1	133	1 RS8E_METKA	Q8TV55 methanopyru
4	34	79.1	475	1 YEB4_HAEIN	P44693 haemophilus
5	34	79.1	497	1 HME5_DROME	P18488 drosophila
6	34	79.1	605	1 ALS_HUMAN	P35858 homo sapien
7	34	79.1	605	1 ALS_PAPHA	O02833 papio hamad
8	33	76.7	99	1 FIXX_RHILP	Q05561 rhizobium 1
9	33	76.7	711	1 RED1_RAT	P51400 rattus norv
10	32	74.4	156	1 SPRT_HAEIN	P44119 haemophilus
11	32	74.4	341	1 YSX3_CABEL	Q10022 caenorhabdi
12	32	74.4	492	1 MOT3_MOUSE	O35308 mus musculu
13	32	74.4	492	1 MOT3_RAT	O70461 rattus norv
14	32	74.4	1196	1 HK2_MOUSE	Q9QZ5 mus musculu
15	32	74.4	1198	1 HK2_HUMAN	Q9H2X6 homo sapien
16	32	74.4	1225	1 NKT5_MOUSE	Q9W30 mus musculu
17	32	74.4	1531	1 NKT5_HUMAN	Q94916 homo sapien
18	32	74.4	1736	1 ZOL_HUMAN	Q07157 homo sapien
19	31	72.1	159	1 MOAE_RHILO	Q984P0 rhizobium 1
20	31	72.1	170	1 RL17_HORVU	P35266 hordeum vul
21	31	72.1	171	1 RL17_MAIZE	O48557 zea mays (m
22	31	72.1	172	1 RL17_HORVU	P35267 hordeum vul
23	31	72.1	183	1 RL17_PODCA	P37380 podocoryne
24	31	72.1	184	1 RL17_HUMAN	P18621 homo sapien
25	31	72.1	184	1 RL17_MOUSE	Q9CPR4 mus musculu
26	31	72.1	184	1 RL17_RAT	P24049 rattus norv
27	31	72.1	186	1 RL17_NEUCR	Q9HE25 neurospora
28	31	72.1	296	1 MYDM_MOUSE	O35682 mus musculu
29	31	72.1	388	1 GCH2_THEMEA	Q9X2E6 thermotoga
30	31	72.1	413	1 COBL_PSEDE	P21921 pseudomonas
31	31	72.1	457	1 TMS5_HUMAN	Q9H3S3 homo sapien
32	31	72.1	674	1 RGS9_HUMAN	O75916 homo sapien
33	31	72.1	811	1 GAG_FOAMV	P14349 human spuma

Q13164 homo sapien
Q13508 penicillium
P27401 simian foam
P23074 simian foam
P26485 azorhizobiu
P27329 lily virus
P36734 human papil
P98153 homo sapien
O52723 rhodobacter
P12807 pichia angu
Q90625 gallus gall
P78574 aspergillus

34 31 72.1 815 1 MK07_HUMAN
35 31 72.1 860 1 AREA_PENRO
36 31 72.1 1157 1 POL_SFV3L
37 31 72.1 1161 1 POL_SFV1
38 30 69.8 97 1 FIXX_AZOCA
39 30 69.8 216 1 VHEL_LVX
40 30 69.8 368 1 VE2_HP45
41 30 69.8 550 1 IDD_HUMAN
42 30 69.8 672 1 TKT_RHOCA
43 30 69.8 692 1 AWO_PICAN
44 30 69.8 706 1 Z151_CHICK
45 30 69.8 749 1 CATA_ASPFU

ALIGNMENTS

RESULT 1

NGP1_HUMAN
ID NGP1_HUMAN STANDARD; PRT; 731 AA.
AC Q13823; Q9BNW7;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Autoantigen NGP-1.
GN NGP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Breast carcinoma;
RX MEDLINE=96419438; PubMed=8822211;
RA Racevskis J., Dill A., Stockert R., Fineberg S.A.;
RT "Cloning of a novel nucleolar guanosine 5'-triphosphate binding
protein autoantigen from a breast tumor.";
RL Cell Growth Differ. 7:271-280(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC -!- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED AT RELATIVELY LOW
LEVELS IN ALL HUMAN TISSUES TESTED, WITH THE HIGHEST LEVEL OF
EXPRESSION IN THE TESTES.
CC -!- SIMILARITY: BELONGS TO THE MMRI/HSR1 FAMILY OF GTP-BINDING
PROTEINS. STRONG, TO YEAST YNR053C.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; L05425; AAC37588.1; -
EMBL; BC000107; AAH00107.1; -
DR SWISS-2DPAGE; Q13823; HUMAN.
DR InterPro; IPR005289; GTP-binding_dom.
DR InterPro; IPR002917; MMR_HSR1.
DR Pfam; PF01926; MMR_HSR1; 1.
DR TIGRfams; TIGR00650; MG442; 1.
KW GTP-binding; Nuclear protein.
FT NP_BIND 317 324 GTP (POTENTIAL).
FT NP_BIND 361 365 GTP (POTENTIAL).
FT DOMAIN 565 571 POLY-GLU.
FT CONFLICT 284 284 S -> I (IN REF. 2).
SQ SEQUENCE 731 AA; 83655 MW; E90C638B58C0886E CRC64;

Query Match 81.4%; Score 35; DB 1; Length 731;

Wed Jul 16 13:41:36 2003

us-09-910-582b-3.rsp

```

RA Malykh A.G., Koonin E.V., Kozhavkin S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RL and monophyly of archaeal methanogens.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
CC -|- SIMILARITY: BELONGS TO THE S8E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE010445; AM02758.1;
CC PROSITE: PS01193; RIBOSOMAL_S8E: 1.
CC RIBOSOMAL PROTEIN; Complete proteome.
SQ SEQUENCE 133 AA; 14876 MW; 2B3470E3FDB87DC0 CRC64;

Query Match 79.1%; Score 34; DB 1; Length 133;
Best Local Similarity 83.3%; Pred. No. 3.4;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRRVRPH 7
DB 15 GRIRPH 20

RESULT 4
YEBA_HAEIN STANDARD; PRT; 475 AA.
ID YEBA_HAEIN AC P44693;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein HI0409.
GN HI0409.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OC NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kervatage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -|- FUNCTION: COULD BE INVOLVED IN CELL WALL DEGRADATION OR
CC FORMATION (BY SIMILARITY).
CC -|- SIMILARITY: STRONG, TO E. COLI YEBA.
CC -|- SIMILARITY: TO STAPHYLOCOCCUS LYSOSTAPHIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32724; AAC22068.1;
CC MEROPS: M37 UPW;
CC TIGR: HI0409;

DR
DR
DR

```

```

RA Best Local Similarity 71.4%; Pred. No. 12;
RA Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
DB 122 HDRIRPH 128

RESULT 2
YRP3_AZOCA STANDARD; PRT; 68 AA.
ID YRP3_AZOCA AC P33988;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in RPON 3' region (ORF3) (Fragment).
OS Azorhizobium caulinodans.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Hypomicrobium group; Azorhizobium.
OC NCBI_TaxID=7;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORS571;
RX MEDLINE=93229807; PubMed=8471796;
RA Stigter J., Schneider M., de Bruijn F.J.;
RA "Azorhizobium caulinodans nitrogen fixation (nif/fix) gene
RT regulation: mutagenesis of the nifA -24/-12 promoter element,
RT characterization of a ntrA(rpoN) gene, and derivation of a model.";
RL Mol. Plant Microbe Interact. 6:238-252(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X69959; CAA49583.1;
CC PIR: S33581;
CC Hypothetical protein.
KW NON_TER 68
FT SEQUENCE 68 AA; 7612 MW; 435C7F141E9C54EB CRC64;

Query Match 79.1%; Score 34; DB 1; Length 68;
Best Local Similarity 85.7%; Pred. No. 1.7;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
DB 38 HGRVRPH 44

RESULT 3
RS8E_METKA STANDARD; PRT; 133 AA.
ID RS8E_METKA AC Q8TV55;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S8e.
GN RPS8E OR RPS8A OR MK1545.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyraceae;
OC Methanopyrus.
OC NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,

```

DR InterPro: IPR002886; Peptidase_M37.
 DR Pfam: PF01551; Peptidase_M37; 1.
 KW Hypothetical protein; Cell wall; Complete proteome.
 SQ SEQUENCE 475 AA; 53255 MW; 143C10F92233939D CRC64;

Query Match 79.1%; Score 34; DB 1; Length 475;
 Best Local Similarity 83.3%; Pred. No. 13;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVRVRPH 7
 II:III
 343 GRIRPH 348

Db

RESULT 5

HMS_DROME STANDARD; PRT; 497 AA.

AC P18488; Q9VFQ1;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Empty spiracles homeotic protein.
 GN EMS OR E4 OR C82988.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Canton-S;
 RX MEDLINE=92289690; PubMed=1376248;
 RA Walldorf U., Gehring W.J.;
 RT "Empty spiracles, a gap gene containing a homeobox involved in
 Drosophila head development.";
 RL EMBO J. 11:2247-2259(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Oregon-R;
 RX MEDLINE=90152339; PubMed=2576012;
 RA Dalton D., Chadwick R., McGinnis W.;
 RT "Expression and embryonic function of empty spiracles: a Drosophila
 homeo box gene with two patterning functions on the
 anterior-posterior axis of the embryo.";
 RL Genes Dev. 3:1940-1956(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brannon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: ACTS AS A HOMEOTIC SELECTOR GENE CONTROLLING ANTENNAL
 CC AND MANDIBULAR SEGMENT IDENTITY.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- DEVELOPMENTAL STAGE: EMS HAS TWO DIFFERENT SPATIAL PATTERNS OF
 CC EXPRESSION DURING EMBRYOGENESIS. THE EMS HEAD-SPECIFIC EXPRESSION
 CC PATTERN INITIATES PRIOR TO CELLULAR BLASTODERM AND CONTINUES ONLY
 CC UNTIL EARLY GERM-BAND EXTENSION. THE EMS METAMERIC EXPRESSION
 CC PATTERN INITIATES AFTER GASTRULATION AND IS EXPRESSED IN THE
 CC LATERAL NEUROBLASTS, IN ECTODERMAL CELLS AT THE ANTERIOR LATERAL
 CC BORDERS OF EACH SEGMENT, AND IN THE FILZKOEPER PRIMORDIA.
 CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF STRAIN CANTON S.
 CC -1- SIMILARITY: BELONGS TO THE EMX HOMEBOX FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X66270; CAA46985.1; -;
 CC EMBL: X51653; CAA35965.1; -;
 CC EMBL: AE003702; AAF54999.1; -;
 CC PIR: A34218; A34218.
 CC PIR: S22708; S22708.
 CC HSP: P06601; IFLJ.
 CC TRANSFAC: T02008; -;
 CC FlyBase: FBgn0000576; ems.
 CC InterPro: IPR001356; Homeobox.
 CC Pfam: PF00046; homeobox; 1.
 CC ProDom: PD000010; Homeobox; 1.
 CC SMART: SM00389; HOX; 1.
 CC PROSITE: PS00027; HOMEBOX_1; 1.
 CC PROSITE: PS00071; HOMEBOX_2; 1.
 CC Homeobox; DNA-binding; Developmental protein; Nuclear protein.
 KW DOMAIN 1 387 PRO-RICH.
 FT DOMAIN 99 362 GLN-RICH.
 FT DOMAIN 334 341 POLY-ALA.
 FT DNA_BIND 391 450 HOMEBOX.
 FT DOMAIN 471 497 ASP/GLU-RICH (ACIDIC).
 FT VARIANT 4 4 M -> T (IN STRAIN OREGON-R).
 FT VARIANT 102 104 MISSING (IN STRAINS BERKELEY AND OREGON-
 FT VARIANT 131 131 H -> Q (IN STRAINS BERKELEY AND OREGON-
 FT VARIANT 158 159 SG -> TR (IN STRAIN OREGON-R).
 FT VARIANT 304 305 EL -> DV (IN STRAIN OREGON-R).
 SQ SEQUENCE 497 AA; 53784 MW; BB7E4548155C0A5C CRC64;

Query Match 79.1%; Score 34; DB 1; Length 497;
 Best Local Similarity 71.4%; Pred. No. 13;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
 III:II
 371 HGRIFPH 377

Db

Wed Jul 16 13:41:36 2003

us-09-910-582b-3.rsp

```

RESULT 6
ALN_HUMAN
ID ALN_HUMAN STANDARD; PRT; 605 AA.
AC P35858;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Insulin-like growth factor binding protein complex acid labile chain
DE precursor (ALS).
GN IGFALS OR ALS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=92357025; PubMed=1379671;
RA Leong S.R., Baxter R.C., Camerato T., Dai J., Wood W.I.;
RT "Structure and functional expression of the acid-labile subunit of
RL the insulin-like growth factor-binding protein complex.";
RL Mol. Endocrinol. 6:870-876(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Frankland J.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE OF 28-35.
MEDLINE=89308584; PubMed=2473065;
RA Baxter R.C., Martin J.L., Beniac V.A.;
RT "High molecular weight insulin-like growth factor binding protein
complex. Purification and properties of the acid-labile subunit from
human serum.";
RT J. Biol. Chem. 264:11843-11848(1989).
RL CC -1- FUNCTION: INVOLVED IN PROTEIN-PROTEIN INTERACTIONS THAT RESULT
IN PROTEIN COMPLEXES, RECEPTOR-LIGAND BINDING OR CELL ADHESION.
CC CC -1- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 KDA WITH
IGF-I OR IGF-II AND IGFBP-3.
CC CC -1- SUBCELLULAR LOCATION: Extracellular.
CC CC -1- TISSUE SPECIFICITY: PLASMA.
CC CC -1- SIMILARITY: CONTAINS 20 LEUCINE-RICH REPEATS (LRR).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M86826; AAA36047.1;
EMBL; AL031724; CAC36078.1;
PIR; A41915; A41915
Genew; HGNC:5468; IGFALS.
MIM; 601489;
InterPro; IPR001611; LRR.
InterPro; IPR000483; LRR_Cterm.
InterPro; IPR000372; LRR_Nterm.
InterPro; IPR003592; LRR_out.
InterPro; IPR003591; LRR_typ.
Pfam; PF00560; LRR; 19.
Pfam; PF01462; LRRNT; 1.
Pfam; PF01463; LRRCT; 1.
PRINTS; PR00019; LEURICRPT.
SMART; SM00370; LRR; 2.
SMART; SM00082; LRRCT; 1.
SMART; SM00013; LRRNT; 1.
SMART; SM00369; LRR_TYP; 11.
KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.
FT SIGNAL 1 27
FT CHAIN 28 605 INSULIN-LIKE GROWTH FACTOR BINDING
FT REPEAT 53 73 PROTEIN COMPLEX ACID LABILE CHAIN.
FT LRR 1.

Query Match 79.1%; Score 34; DB 1; Length 605;
Best Local Similarity 83.3%; Pred. No. 16;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRVRPH 7
DB 399 GRIRPH 404

RESULT 7
ALN_PAPHA
ID ALN_PAPHA STANDARD; PRT; 605 AA.
AC 002833;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Insulin-like growth factor binding protein complex acid labile chain
DE precursor (ALS).
GN IGFALS OR ALS.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
OX NCBI_TaxID=9557;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=97040714; PubMed=8886027;
RA Delhanty P., Baxter R.C.;
RT "The cloning and expression of the baboon acid-labile subunit of the
RT insulin-like growth factor binding protein complex.";
RL Biochem. Biophys. Res. Commun. 227:897-902(1996).
CC CC -1- FUNCTION: INVOLVED IN PROTEIN-PROTEIN INTERACTIONS THAT RESULT
IN PROTEIN COMPLEXES, RECEPTOR-LIGAND BINDING OR CELL ADHESION.
CC CC -1- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 KDA WITH
IGF-I OR IGF-II AND IGFBP-3 (BY SIMILARITY).
CC CC -1- SUBCELLULAR LOCATION: Extracellular.
CC CC -1- SIMILARITY: CONTAINS 20 LEUCINE-RICH REPEATS (LRR).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----

```

```
-----NOT_ANNOTATED_CDS.-----
CC EMBL: S83462; NOT_ANNOTATED_CDS.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR000359; LRR_Out.
DR InterPro: IPR003591; LRR_Typ.
DR Pfam: PF00560; LRR; 19.
DR Pfam: PF01462; LRRNT; 1.
DR Pfam: PF01463; LRRCT; 1.
DR PRINTS: PR00019; LEURICHRPT.
DR SMART: SM00370; LRR; 2.
DR SMART: SM00082; LRRCT; 1.
DR SMART: SM00013; LRRNT; 1.
DR SMART: SM00369; LRR_Typ; 11.
KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.
FT SIGNAL 1 27 BY SIMILARITY.
FT CHAIN 28 605 INSULIN-LIKE GROWTH FACTOR BINDING
PROTEIN COMPLEX ACID LABILE CHAIN.
FT REPEAT 53 73 LRR 1.
FT REPEAT 74 96 LRR 2.
FT REPEAT 98 120 LRR 3.
FT REPEAT 121 144 LRR 4.
FT REPEAT 146 168 LRR 5.
FT REPEAT 169 192 LRR 6.
FT REPEAT 194 216 LRR 7.
FT REPEAT 217 240 LRR 8.
FT REPEAT 242 264 LRR 9.
FT REPEAT 265 288 LRR 10.
FT REPEAT 289 312 LRR 11.
FT REPEAT 313 336 LRR 12.
FT REPEAT 338 360 LRR 13.
FT REPEAT 361 384 LRR 14.
FT REPEAT 386 408 LRR 15.
FT REPEAT 409 432 LRR 16.
FT REPEAT 433 456 LRR 17.
FT REPEAT 458 480 LRR 18.
FT REPEAT 482 504 LRR 19.
FT REPEAT 505 530 LRR 20.
FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 580 580 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 605 AA; 66110 MW; 9D71994625F23652 CRC64;

Query Match 79.1%; Score 34; DB 1; Length 605;
Best Local Similarity 83.3%; Pred. No. 16;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRVRPH 7
Db 399 GRIRPH 404

RESULT 8
FIXX_RHILP STANDARD; PRT; 99 AA.
ID FIXX_RHILP
AC Q05561;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ferredoxin-like protein.
GN FIXX.
OS Rhizobium leguminosarum (biovar phaseoli).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=385;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93379046; PubMed=8369342;
RA Michiels J., Vanderleyden J.;

-----
RT RT Cloning and sequence of the Rhizobium leguminosarum biovar phaseoli
fixa gene".
RL Blochim. Biophys. Acta 1144:232-233(1993).
CC -1- FUNCTION: COULD BE A 3FE-4S CLUSTER-CONTAINING PROTEIN.
CC -1- SIMILARITY: TO FERREDOXIN FROM PSEUDOMONAS PUTIDA AND CLOSTRIDIUM
TARTARIVORUM, FERREDOXIN I FROM ACETOBACTER VINELANDII, FERREDOXIN
II FROM DESULFOVIBRIO DESULFURICANS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L11082; AAA26270.1;
DR InterPro: IPR001450; 4Fe4S-ferredoxin.
DR PROSITE: PS00198; 4FE4S-FERREDOXIN; 1.
KW Electron transport; Iron-sulfur; Nitrogen fixation.
SQ SEQUENCE 99 AA; 11114 MW; 2B46B5D60613F0DF CRC64;

Query Match 76.7%; Score 33; DB 1; Length 99;
Best Local Similarity 85.7%; Pred. No. 4;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
Db 27 HIRVRPH 33

RESULT 9
RED1_RAT STANDARD; PRT; 711 AA.
ID RED1_RAT
AC P51400;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Double-stranded RNA-specific editase 1 (EC 3.5.-.-) (dsRNA adenosine
deaminase) (RNA editing enzyme 1).
DE ADAR1 OR RED1.
GN Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96158881; PubMed=8559253;
RA Melcher T., Maas S., Herb A., Sprengel R., Seeburg P.H., Higuchi M.;
RT "A mammalian RNA editing enzyme.";
RL Nature 379:460-464(1996).
CC -1- FUNCTION: EDITING OF THE MESSENGER RNAs FOR GLUTAMATE RECEPTOR
(GLUR) SUBUNITS BY SITE-SELECTIVE ADENOSINE DEAMINATION. EDITS
BOTH THE GLUR-B O/R AND R/G SITES EFFICIENTLY BUT CONVERTS THE
ADENOSINE IN HOTSPOT1 MUCH LESS EFFICIENTLY.
CC -1- TISSUE SPECIFICITY: BRAIN AND PERIPHERAL TISSUES.
CC -1- SIMILARITY: BELONGS TO THE ADENOSINE DEAMINASE EDITASE FAMILY.
CC -1- SIMILARITY: CONTAINS 2 DBM (DOUBLE-STRANDED RNA-BINDING) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U43534; AAA96755.1;
DR HSPSP; Q91836; IDI2
DR InterPro: IPR002466; A_deamin.
DR InterPro: IPR001159; DS_RBD.
DR Pfam: PF00035; dsrm; 2.
DR Pfam: PF02137; A_deamin; 1.
```

Wed Jul 16 13:41:36 2003

us-09-910-582b-3.rsp

SMART: SM00358; DSRM; 2.
DR PROSITE: PS0141; A-DEAMIN-EDITASE; 1.
DR PROSITE: PS0137; DS-RBD; 2.
KW mRNA processing; Hydrolase; zinc; RNA-binding; Repeat.
FT DOMAIN 76 146 DRBM 1.
FT METAL 230 300 DRBM 2.
FT METAL 394 394 ZINC (BY SIMILARITY).
FT ACT_SITE 396 396 BY SIMILARITY.
FT METAL 451 451 ZINC (BY SIMILARITY).
FT METAL 526 526 ZINC (BY SIMILARITY).
SQ SEQUENCE 711 AA; 77924 MW; 5A25C4B202530C54 CRC64;

Query Match 76.7%; Score 33; DB 1; Length 711;
Best Local Similarity 71.4%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 HGRVPRH 7
Db 650 HGKVPFH 656

RESULT 10
ID SPRT_HAEIN STANDARD; PRT; 156 AA.

AC P44119;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein sppt homolog.
GN SPRT OR H1173
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
CC SEQUENCE FROM N.A.
DR STRAIN-Rd / KW20 / ATCC 51907;
DR MEDLINE-95350630; PubMed-7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kierulff A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA Kierulff A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
CC "Whole-genome random sequencing and assembly of Haemophilus influenzae
Rd.";
CC Science 269:496-512(1995).
CC -1- SIMILARITY: STRONG, TO E-COLI SPRT.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

EMBL: U32797; AAC22826.1;
TIGR: H1173;
DR InterPro: IPR000130; Zn_MTPeptidase.
DR PROSITE: PS0142; ZINC_PROTEASE; UNKNOWN_1.
KW Zinc; Complete proteome.
FT METAL 65 65 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 66 66 BY SIMILARITY.
FT METAL 69 69 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 156 AA; 18684 MW; BBEF66F0585C465A CRC64;

Query Match 74.4%; Score 32; DB 1; Length 156;
Best Local Similarity 83.3%; Pred. No. 10;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GRVPRH 7
Db 77 GRVPRH 82

RESULT 11
ID YSX3_CAEEL STANDARD; PRT; 341 AA.

AC Q10022;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 39.0 kDa protein T28D9.3 in chromosome II.
GN T28D9.3
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
CC SEQUENCE FROM N.A.
DR STRAIN-Bristol N2;
RA Waterston R.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: WEAK, TO YEAST D9719.9.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

EMBL: U28738; AAA68312.1;
DR Wormpep; T28D9.3; CE02068
DR InterPro: IPR000326; PA_PTPase.
DR Pfam: PF01569; PAP2; 1.
DR SMART: SM00014; acidppc; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 30 50 POTENTIAL.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 223 243 POTENTIAL.
FT TRANSMEM 257 277 POTENTIAL.
SQ SEQUENCE 341 AA; 39028 MW; DF74E39BC3E2DA8E CRC64;

Query Match 74.4%; Score 32; DB 1; Length 341;
Best Local Similarity 83.3%; Pred. No. 23;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GRVPRH 7
Db 148 GRLRPH 153

RESULT 12
ID MOT3_MOUSE STANDARD; PRT; 492 AA.

AC O35308;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Monocarboxylate transporter 3 (MCT 3) (Proton-coupled monocarboxylate
transporter 3).
GN SLC16A8 OR MCT3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
CC SEQUENCE FROM N.A.

RC STRAIN-C3H, and 129/Sv;
 RX MEDLINE=21185340; PubMed=11287345;
 RA Philp N.J., Yoon H., Lombardi L.;
 RT "Mouse MCT3 gene is expressed preferentially in retinal pigment and
 choroid plexus epithelia.";
 RL Am. J. Physiol. 280:C1319-C1326(2001).
 CC -1- FUNCTION: PROTON-LINKED MONOCARBOXYLATE TRANSPORTER. CATALYZES THE
 CC RAPID TRANSPORT ACROSS THE PLASMA MEMBRANE OF MANY
 CC MONOCARBOXYLATES SUCH AS LACTATE, PYRUVATE, BRANCHED-CHAIN OXO
 CC ACIDS DERIVED FROM LEUCINE, VALINE AND ISOLEUCINE, AND THE KETONE
 CC BODIES ACETOACETATE, BETA-HYDROXYBUTYRATE AND ACETATE (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Plasma membrane.
 CC -1- TISSUE SPECIFICITY: RETINAL PIGMENT EPITHELIUM.
 CC -1- SIMILARITY: BELONGS TO THE SLC16 FAMILY OF TRANSPORTERS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF019111; AAB70592.2; -;
 CC EMBL; AF178956; AAF45042.1; -;
 CC MGD; MG1:1929519; SLC16a8.
 CC InterPro; IPR004743; Mcarb_transport.
 CC TIGRFAMs; TIGR00892; 2A0113; 1.
 CC Transports; Symport; Transmembrane; Multigene family.
 CC DOMAIN 1 14 CYTOPLASMIC (POTENTIAL).
 CC TRANSSEM 15 35 POTENTIAL.
 CC DOMAIN 36 58 EXTRACELLULAR (POTENTIAL).
 CC TRANSSEM 59 79 POTENTIAL.
 CC DOMAIN 80 85 CYTOPLASMIC (POTENTIAL).
 CC TRANSSEM 86 106 POTENTIAL.
 CC DOMAIN 107 115 EXTRACELLULAR (POTENTIAL).
 CC TRANSSEM 116 136 POTENTIAL.
 CC DOMAIN 137 146 CYTOPLASMIC (POTENTIAL).
 CC TRANSSEM 147 167 POTENTIAL.
 CC DOMAIN 168 172 EXTRACELLULAR (POTENTIAL).
 CC TRANSSEM 173 193 POTENTIAL.
 CC DOMAIN 194 228 CYTOPLASMIC (POTENTIAL).
 CC TRANSSEM 229 249 POTENTIAL.
 CC DOMAIN 250 257 EXTRACELLULAR (POTENTIAL).
 CC TRANSSEM 258 278 POTENTIAL.
 CC DOMAIN 279 293 CYTOPLASMIC (POTENTIAL).
 CC TRANSSEM 294 314 POTENTIAL.
 CC DOMAIN 315 318 EXTRACELLULAR (POTENTIAL).
 CC TRANSSEM 319 339 POTENTIAL.
 CC DOMAIN 340 352 CYTOPLASMIC (POTENTIAL).
 CC TRANSSEM 353 373 POTENTIAL.
 CC DOMAIN 374 386 EXTRACELLULAR (POTENTIAL).
 CC TRANSSEM 387 407 POTENTIAL.
 CC DOMAIN 408 492 CYTOPLASMIC (POTENTIAL).
 CC SEQUENCE 492 AA; 51556 MW; 2B19E362B96D9ACC CRC64;
 Query Match 74.4% Score 32; DB 1; Length 492;
 Best Local Similarity 83.3%; Pred. No. 34;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 GRVRPH 7
 Db 289 GRLRPH 294
 ||:||||
 STANDARD; PRT; 492 AA.
 RESULT 13
 MOT3_RAT
 ID MOT3_RAT
 AC O70461;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Monocarboxylate transporter 3 (MCT 3).
 GN SLC16A8 OR MCT3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Retina;
 RX MEDLINE=99020931; PubMed=9841555;
 RA Philp N.J., Yoon H., Grollman E.P.;
 RT "Monocarboxylate transporter MCT1 is located in the apical membrane
 and MCT3 in the basal membrane of rat RPE.";
 RL Am. J. Physiol. 274:R1824-R1828(1998).
 CC -1- FUNCTION: PROTON-LINKED MONOCARBOXYLATE TRANSPORTER. CATALYZES THE
 CC RAPID TRANSPORT ACROSS THE PLASMA MEMBRANE OF MANY
 CC MONOCARBOXYLATES SUCH AS LACTATE, PYRUVATE, BRANCHED-CHAIN OXO
 CC ACIDS DERIVED FROM LEUCINE, VALINE AND ISOLEUCINE, AND THE KETONE
 CC BODIES ACETOACETATE, BETA-HYDROXYBUTYRATE AND ACETATE (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Plasma membrane.
 CC -1- TISSUE SPECIFICITY: RETINAL PIGMENT EPITHELIUM. RESTRICTED TO THE
 CC BASOLATERAL MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE SLC16 FAMILY OF TRANSPORTERS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF059258; AAC18120.1; -;
 CC InterPro; IPR004743; Mcarb_transport.
 CC TIGRFAMs; TIGR00892; 2A0113; 1.
 CC Transports; Symport; Transmembrane; Multigene family.
 CC DOMAIN 1 14 CYTOPLASMIC (POTENTIAL).
 CC TRANSSEM 15 35 POTENTIAL.
 CC DOMAIN 36 58 EXTRACELLULAR (POTENTIAL).
 CC TRANSSEM 59 79 POTENTIAL.
 CC DOMAIN 80 85 CYTOPLASMIC (POTENTIAL).
 CC TRANSSEM 86 106 POTENTIAL.
 CC DOMAIN 107 115 EXTRACELLULAR (POTENTIAL).
 CC TRANSSEM 116 136 POTENTIAL.
 CC DOMAIN 137 147 CYTOPLASMIC (POTENTIAL).
 CC TRANSSEM 148 168 POTENTIAL.
 CC DOMAIN 169 172 EXTRACELLULAR (POTENTIAL).
 CC TRANSSEM 173 193 POTENTIAL.
 CC DOMAIN 194 230 CYTOPLASMIC (POTENTIAL).
 CC TRANSSEM 231 251 POTENTIAL.
 CC DOMAIN 252 257 EXTRACELLULAR (POTENTIAL).
 CC TRANSSEM 258 278 POTENTIAL.
 CC DOMAIN 279 293 CYTOPLASMIC (POTENTIAL).
 CC TRANSSEM 294 314 POTENTIAL.
 CC DOMAIN 315 318 EXTRACELLULAR (POTENTIAL).
 CC TRANSSEM 319 339 POTENTIAL.
 CC DOMAIN 340 352 CYTOPLASMIC (POTENTIAL).
 CC TRANSSEM 353 373 POTENTIAL.
 CC DOMAIN 374 386 EXTRACELLULAR (POTENTIAL).
 CC TRANSSEM 387 407 POTENTIAL.
 CC DOMAIN 408 492 CYTOPLASMIC (POTENTIAL).
 CC SEQUENCE 492 AA; 51590 MW; 5CCC44DC38B35FA0 CRC64;
 Query Match 74.4% Score 32; DB 1; Length 492;
 Best Local Similarity 83.3%; Pred. No. 34;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 GRVRPH 7
 Db 289 GRLRPH 294
 ||:||||

```

RESULT 14
HIK2_MOUSE STANDARD; PRT; 1196 AA.
AC Q902R5; O88905; Q9EQ22;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Homeodomain-interacting protein kinase 2 (EC 2.7.1.-) (Nuclear body
DE associated kinase 1) (Sialophorin tail associated nuclear
DE serine/threonine kinase).
GN HIKP2 OR NBAK1 OR STANK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN-BALB/c;
RC MEDLINE=98421509; PubMed=9748262;
-RX Kim Y.H., Choi C.Y., Lee S.J., Conti M.A., Kim Y.;
RT "Homeodomain-interacting protein kinases, a novel family of co-
RT repressors for homeodomain transcription factors.";
RL J. Biol. Chem. 273:25875-25879(1998).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORM 2).
RA Sather S.L., Johnson N.L., Johnson G.L.;
RA "Protein kinases associated with PML/CBP nuclear bodies and
RA filamentous threads regulate transcription and inhibit cell growth.";
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RL [3]
RN SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN-BALB/c;
RC MEDLINE=20532484; PubMed=11078605;
-RX Wang W., Link V., Green J.M.;
RT "Identification and cloning of a CD43-associated serine/threonine
RT kinase.";
RL Cell. Immunol. 205:34-39(2000).
RN [4]
RN SEQUENCE OF 8-1196 FROM N.A. (ISOFORM 2).
RX MEDLINE=21167386; PubMed=11267674;
RA Wang Y., Hofmann T.G., Runkel L., Haaf T., Schaller H., Debatin K.-M.,
RA Hug H.;
RT "Isolation and characterization of cDNAs for the protein kinase
RT HIKP2";
RL Biochim. Biophys. Acta 1518:168-172(2001).
RN [5]
RN SUMOYLATION.
RX MEDLINE=20006237; PubMed=10535925;
RA Kim Y.H., Choi C.Y., Kim Y.;
RT "Covalent modification of the homeodomain-interacting protein kinase 2
RT (HIKP2) by the ubiquitin-like protein SUMO-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:12350-12355(1999).
CC -1- FUNCTION: PHOSPHORYLATES HOMEODOMAIN TRANSCRIPTION FACTORS. MAY
CC PLAY A ROLE AS A CO-REPRESSOR FOR HOMEODOMAIN TRANSCRIPTION
CC FACTORS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- PTM: Sumoylated. When conjugated it is directed to nuclear
CC speckles.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. HIPK
CC SUBFAMILY.
CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL: AF077659; AAC63011.1; ALT_FRAME.
DR EMBL: AF170301; AAD52566.1;
DR EMBL: AF273680; AAG02078.1;
DR EMBL: AF208292; AAG41237.1;
DR HSSP: P24941; 1CRP.
DR MGD: MGI:1314872; Hipk2.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00219; TyrcK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR transferase; Serine/threonine-protein kinase; ATP-binding;
KW Nuclear protein; Alternative splicing; Ub1 conjugation.
FT DOMAIN 199 527 PROTEIN KINASE.
FT NP_BIND 205 213 ATP (PROBABLE).
FT BINDING 228 228 ATP (PROBABLE).
FT ACT_SITE 324 324 POLY-ALA.
FT DOMAIN 1089 1092 MISSING (IN ISOFORM 2).
FT VARSPIC 595 621 I -> T (IN REF. 1).
FT CONFLICT 460 460 V -> G (IN REF. 1).
FT CONFLICT 479 479 MISSING (IN REF. 4).
FT CONFLICT 705 705 A -> T (IN REF. 1).
FT CONFLICT 719 719 A -> R (IN REF. 1).
FT CONFLICT 1120 1120
FT SEQUENCE 1196 AA; 130498 MW; 5C863BE377F3AAEF CRC64;
SQ

```

```

Query Match 74.48; Score 32; DB 1; Length 1196;
Best Local Similarity 71.48; Pred.No. 84;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 HGRVRPH 7
DB 1081 HGTVRPH 1087

```

```

RESULT 15
HIK2_HUMAN STANDARD; PRT; 1198 AA.
ID HIK2_HUMAN STANDARD; PRT; 1198 AA.
AC Q9H2X6; Q9H2Y1; Q8WWI4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Homeodomain-interacting protein kinase 2 (EC 2.7.1.-).
GN HIPK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE-Testis, and Liver;
RX MEDLINE=21167386; PubMed=11267674;
RA Wang Y., Hofmann T.G., Runkel L., Haaf T., Schaller H., Debatin K.-M.,
RA Hug H.;
RT "Isolation and characterization of cDNAs for the protein kinase
RT HIPK2";
RL Biochim. Biophys. Acta 1518:168-172(2001).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE-Frontal cortex;
RA Stukart G.C., Dias-Neto E.;
RT "Sequencing of hHIPK2, a human homolog of mouse homeodomain
RT interacting protein kinase 2.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE OF 8-1198 FROM N.A. (ISOFORM 2).
RA Pierantoni G.M., Benvenuto G., Chiariotti L., Fusco A.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [4]

```


RP HIPK2-TRADD BINDING.
RX MEDLINE=20490030; PubMed=11032752;
RA Li X., Wang Y., Debatin K.-M., Hug H.:
RT "The serine/threonine kinase HIPK2 interacts with TRADD, but not with
CD95 or TNF-R1 in 293T cells.",
RL Biochem. Biophys. Res. Commun. 277:513-517(2000).
CC -!- FUNCTION: PHOSPHORYLATES HOMEODOMAIN TRANSCRIPTION FACTORS. MAY
CC PLAY A ROLE AS A CO-REPRESSOR FOR HOMEODOMAIN TRANSCRIPTION
CC FACTORS.
CC -!- SUBUNIT: INTERACTS WITH TRADD.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; MAY BE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN NEURONAL TISSUES, HEART
CC AND KIDNEY. WEAKLY EXPRESSED IN AN UBIQUITOUS WAY.
CC -!- PTM: Sumoylated. When conjugated it is directed to nuclear
CC speckles (by similarity).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. HIPK
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF208291; AAG41236.1; -;
DR EMBL; AF326592; AAL37371.1; -;
DR EMBL; AF207702; AAG35710.1; -;
DR Genew; HGNC:14402; HIPK2.
DR MIM; 606868; -;
DR HSSP; P24941; 1A01.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TyrKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Nuclear protein; Alternative splicing
FT DOMAIN 199 527 PROTEIN KINASE.
FT NP_BIND 205 213 ATP (PROBABLE).
FT BINDING 228 228 ATP (PROBABLE).
FT ACT_SITE 324 324 PROBABLE.
FT DOMAIN 1088 1094 POLY-ALA.
FT VARSPPLIC 595 621 MISSING (IN ISOFORM 3).
FT VARSPPLIC 808 907 MISSING (IN ISOFORM 2).
FT VARSPPLIC 989 1018 VNTSHSSYKSKSSNVSTSGHSGSSS -> GNLGPGCQ
FT GRNLSLESGFPAPFLLEMLLYGS (IN ISOFORM 2).
FT VARSPPLIC 1019 1198 MISSING (IN ISOFORM 2).
FT CONFLICT 33 33 I -> V (IN REF. 1).
FT CONFLICT 59 59 L -> P (IN REF. 1).
FT CONFLICT 64 64 T -> S (IN REF. 1).
FT CONFLICT 169 169 S -> F (IN REF. 3).
FT CONFLICT 187 187 V -> S (IN REF. 3).
FT CONFLICT 202 202 L -> S (IN REF. 3).
FT CONFLICT 233 233 H -> R (IN REF. 1).
FT CONFLICT 471 471 N -> I (IN REF. 2).
FT CONFLICT 669 669 P -> S (IN REF. 3).
FT CONFLICT 711 711 T -> N (IN REF. 3).
FT CONFLICT 717 719 PPA -> SPT (IN REF. 3).
FT CONFLICT 724 724 T -> D (IN REF. 3).
SQ SEQUENCE 1198 AA; 130965 MW; 6022D5710E8D2D93 CRC64;

Query Match 74.4%; Score 32; DB 1; Length 1198;
Best Local Similarity 71.4%; Pred. No. 85;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Caps 0;

QY 1 HGRVRPH 7
|||
Db 1080 HGTVHPH 1086

Search completed: July 16, 2003, 13:13:12
Job time : 5.78049 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:07:18 ; Search time 18.9512 Seconds
(without alignments)
76.107 Million cell updates/sec

Title: US-09-910-582B-3
Perfect score: 43
Sequence: 1 HGRVRPH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriaph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	43	100.0	118	10	O22457	O22457 chlamydomon
2	39	90.7	445	3	Q96VB4	Q96VB4 alternaria
3	37	86.0	162	16	Q9KUP4	Q9KUP4 vibrio chol
4	35	81.4	64	10	Q941S8	Q941S8 oryza sativ
5	35	81.4	217	16	Q9FCA2	Q9FCA2 streptomyce
6	35	81.4	229	16	Q9RWD4	Q9RWD4 deinococcus
7	35	81.4	274	10	Q94AY2	Q94AY2 arabidopsis
8	35	81.4	363	10	Q8VXK8	Q8VXK8 ginkgo billo
9	35	81.4	363	12	Q9QTE4	Q9QTE4 marek's dis
10	35	81.4	368	4	Q8WZ58	Q8WZ58 homo sapien
11	35	81.4	404	10	Q8S6L5	Q8S6L5 oryza sativ
12	35	81.4	437	12	Q9E942	Q9E942 gallid herp
13	35	81.4	438	10	Q9FTM3	Q9FTM3 oryza sativ
14	35	81.4	442	16	Q69864	Q69864 streptomyce
15	35	81.4	534	13	Q9DE31	Q9DE31 xenopus lae
16	35	81.4	542	10	Q8RXU2	Q8RXU2 arabidopsis

17	35	81.4	1024	10	O23517	O23517 arabidopsis
18	34	79.1	133	17	O8TV55	O8TV55 methanopyru
19	34	79.1	250	16	Q92BH3	Q92BH3 streptomyce
20	34	79.1	276	5	Q17526	Q17526 caenorhabdi
21	34	79.1	277	3	Q94208	Q94208 penicillium
22	34	79.1	355	16	Q9HZ89	Q9HZ89 pseudomonas
23	34	79.1	422	17	Q974N2	Q974N2 sulfolobus
24	34	79.1	605	4	Q8TAV0	Q8TAV0 homo sapien
25	34	79.1	1156	15	Q93209	Q93209 feline foam
26	33	76.7	206	16	Q91681	Q91681 pseudomonas
27	33	76.7	341	5	Q8SSI3	Q8SSI3 encephalito
28	33	76.7	395	12	Q914X5	Q914X5 newcastle d
29	33	76.7	423	2	Q9EXL8	Q9EXL8 ralstonia s
30	33	76.7	677	11	Q91ZS7	Q91ZS7 mus musculu
31	33	76.7	687	11	Q91ZS6	Q91ZS6 mus musculu
32	33	76.7	701	11	Q99MU8	Q99MU8 mus musculu
33	33	76.7	701	11	Q91ZS9	Q91ZS9 mus musculu
34	33	76.7	711	11	Q91ZS8	Q91ZS8 mus musculu
35	32.5	75.6	352	16	Q92VI9	Q92VI9 rhizobium m
36	32	74.4	40	12	Q91FJ8	Q91FJ8 chilo iride
37	32	74.4	123	17	Q9YD17	Q9YD17 aeropyrum p
38	32	74.4	165	16	Q9CM18	Q9CM18 pasteurella
39	32	74.4	187	6	Q95L84	Q95L84 oryctolagus
40	32	74.4	175	17	Q80UG5	Q80UG5 pyrococcus
41	32	74.4	233	4	Q99782	Q99782 homo sapien
42	32	74.4	234	2	Q937N4	Q937N4 photorhabdu
43	32	74.4	246	5	Q9V577	Q9V577 drosophila
44	32	74.4	249	16	Q9K9Y9	Q9K9Y9 bacillus ha
45	32	74.4	260	6	Q9GM03	Q9GM03 macaca fasc

ALIGNMENTS

RESULT 1

O22457 ID O22457 PRELIMINARY; PRT; 118 AA.
AC O22457;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Low-carbon dioxide inducible protein.
GN LCI3.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_taxid=3055;
RN [1]

RP SEQUENCE FROM N.A.
RA Somanchi A., Handley E.R., Moroney J.V.;
RT "Chlamydomonas reinhardtii cDNAs upregulated in low-carbon dioxide
conditions: expression and analyses."
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF015661; AAB56991.1; -;
SQ SEQUENCE 118 AA; 12896 MW; 1AC8D6E847957CC7 CRC64;

Query Match 100.0%; Score 43; DB 10; Length 118;

Best Local Similarity 100.0%; Pred. No. 0.72; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

Qy

1 HGRVRPH 7
|||||||

34 HGRVRPH 40

RESULT 2

Q96VB4 ID Q96VB4 PRELIMINARY; PRT; 445 AA.
AC Q96VB4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Atr-1.

AFTR-1.
OS Alternaria alternata.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
OC Pleosporales; Pleosporaceae; mitosporic Pleosporaceae; Alternaria.
OX NCBI_TaxID=5599;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=NAF8;
RA Hatta R., Ito K., Hosaki Y., Tanaka T., Tanaka A., Yamamoto M.,
RA Akimitsu K., Tsuge T.;
RT "A conditionally dispensable chromosome controls host-specific
pathogenicity in the fungal plant pathogen Alternaria alternata."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -|- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -|- SIMILARITY: CONTAINS 1 FUNGAL-TYPE BINUCLEAR CLUSTER DOMAIN.
DR EMBL: AB070712; BAB69077.1; -;
DR InterPro: IPR001138; Fungi_Trn.
DR Pfam: PF00172; Zn_clus_1.
DR PROSITE: PS00463; ZN2_CV6_FUNGAL_1; UNKNOWN_1.
DR PROSITE: PS50048; ZN2_CV6_FUNGAL_2; 1.
DR DNA-binding: Metal-binding; Nuclear protein; Transcription regulation;
KW Zinc.
SQ SEQUENCE 445 AA; 48818 MW; 84CC19414E557860 CRC64;
Query Match 90.7%; Score 39; DB 3; Length 445;
Best Local Similarity 85.7%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 HGRVRPH 7
Db 207 HGRTRPH 213
RESULT 3
Q9KUP4 PRELIMINARY; PRT; 162 AA.
AC Q9KUP4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE SpT protein, putative.
GN VC0471.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RC MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae".
RL Nature 406:477-483(2000).
DR EMBL: AE004133; AAF93844.1; -;
DR TIGR: VC0471; -;
DR InterPro: IPR00130; Zn_MTpeptdse.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Complete proteome..
SQ SEQUENCE 162 AA; 18933 MW; E189C3D26BDB7E7A CRC64;
Query Match 86.08; Score 37; DB 16; Length 162;
Best Local Similarity 85.7%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 HGRVRPH 7
Db 85 YGRVRPH 91

RESULT 4
Q941S8 PRELIMINARY; PRT; 64 AA.
AC Q941S8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE P0005H10.20 protein.
GN P0005H10.20.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0005H10.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP004127; BAB64283.1; -;
SQ SEQUENCE 64 AA; 6925 MW; 585C440344A722BC CRC64;
Query Match 81.4%; Score 35; DB 10; Length 64;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GVRVRPH 7
Db 8 GVRVRPH 13
RESULT 5
Q9FCA2 PRELIMINARY; PRT; 217 AA.
AC Q9FCA2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Polypeptide deformylase (EC 3.5.1.31) (PDF) (Formylmethionine
deformylase).
GN DEF OR SC01211 OR 2SCG58.11C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Saunders D.C., Harris D.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RN SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,

RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RL coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 CC -1- FUNCTION: REMOVES THE FORMYL GROUP FROM THE N-TERMINAL MET OF
 CC NEWLY SYNTHESIZED PROTEINS (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: N-FORMYL-L-METHIONINE + H(2)O -> FORMATE + L-
 CC METHIONINE.
 CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.
 DE EMBL: AL391017; CAC01493.1; -;
 DR HSP: P27251; 2DEF.
 DR InterPro: IPR000181; Pep_deformylase.
 DR Pfam: PF01327; Pep_deformylase; 1.
 DR PRINTS: PR01576; PDEFORMLASE.
 DR ProDom: PD003844; Pep_deformylase; 1.
 DR TIGRFAMs: TIGR00079; pep_deformyl; 1.
 KW Hydrolyase; Protein biosynthesis; Zinc.
 SQ SEQUENCE 217 AA; 23770 MW; 6B6E087CBE79B544 CRC64;

Query Match 81.4%; Score 35; DB 16; Length 217;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGRVRP 6
 DB 48 HGRVRP 53

RESULT 6

Q9RWD4 PRELIMINARY; PRT; 229 AA.
 AC Q9RWD4;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Transcriptional regulator, TETR family.
 GN DR0735.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-R1.
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans R1.";
 RL Science 286:1571-1577(1999).
 CC -1- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 DE EMBL: AE001929; AAF10313.1; -;
 DR TIGR: DR0735; -;
 DR InterPro: IPR001647; HTH_Tetr.
 DR Pfam: PF00440; tetr; 1.
 DR PRINTS: PR00455; HTHPETR.
 KW DNA-binding; Transcription regulation; Complete proteome.
 SQ SEQUENCE 229 AA; 25254 MW; 122F571EF81D6024 CRC64;

Query Match 81.4%; Score 35; DB 16; Length 229;

Best Local Similarity 100.0%; Pred. No. 44;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GRVRPH 7
 DB 148 GRVRPH 153

RESULT 7

Q94AY2 PRELIMINARY; PRT; 274 AA.
 AC Q94AY2;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE AT1202080/F9H16.14.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Koesema E., Chen H., Cheuk R., Kim C.J., Meyers M.C., Shinn P.,
 RA Banh J., Bowser L., Carninci P., Dale J.M., Gibson H.A.,
 RA Goldsmith A.D., Hayashizaki Y., Ishida J., Jiang P.X., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J.,
 RA Liu S.X., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,
 RA Davis R.W., Theologis A., Ecker J.R.;
 RT "Arabidopsis cDNA clones.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY045619; AAK73977.1; -;
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00076; rrm; 1.
 DR PROSITE: PS0102; RRM; 1.
 DR PROSITE: PS00030; RRM_RNP_1; UNKNOWN_1.
 SQ SEQUENCE 274 AA; 30036 MW; 17A9AA3C8B4E0050 CRC64;

Query Match 81.4%; Score 35; DB 10; Length 274;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRVRPH 7
 DB 118 GRVRPH 123

RESULT 8

Q8VXK8 PRELIMINARY; PRT; 363 AA.
 AC Q8VXK8;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Phosphoenolpyruvate carboxylase (EC 4.1.1.31) (Fragment).
 GN PEPC.
 OS Ginkgo biloba (Ginkgo).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Ginkgoophyta; Ginkgoales; Ginkgoaceae; Ginkgo.
 OX NCBI_TaxID=3311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gehrig H.H., Virgo A., Cushman J.C., Cushman M., Winter K.;
 RT "Phylogenetic relationship inside modern Gymnosperms: Characterization
 RT of PEPC sequences for evolution studies.";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ312614; CAC84914.1; -;
 DR InterPro: IPR001449; PEPCase.
 DR Pfam: PF00311; PEPCase; 1.
 KW Pyruvate; Lyase.
 FT NON_TER 1 1

Wed Jul 16 13:41:36 2003

us-09-910-582b-3.rspt

```

FT NON_TER 363 363
SQ SEQUENCE 363 AA; 812C5196FB87C391 CRC64;

Query Match 81.4%; Score 35; DB 10; Length 363;
Best Local Similarity 71.4%; Pred. No. 68;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGRVRPH 7
DB 317 HGNLRPH 323

RESULT 9
QOQTEA PRELIMINARY; PRT; 363 AA.
AC QOQTE4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE ORF 4.
OS Marek's disease virus serotype 2 MDV2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=36353;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HPRS24;
RA Jang H., Cai J., Izumiya Y., Murakami Y., Mochizuki M., Song C.,
RA Lee Y., Kai C., Takahashi E., Mikami T.;
RT "The complete DNA sequence and transcription map of the unique long
RT genome region of Marek's disease virus type 2.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB024414; BAA82893.1; -. 53786AB51FF690BA CRC64;
SQ SEQUENCE 363 AA; 40613 MW; 53786AB51FF690BA CRC64;

Query Match 81.4%; Score 35; DB 12; Length 363;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRVRPH 7
DB 152 GRVRPH 157

RESULT 10
Q8WZ58 PRELIMINARY; PRT; 368 AA.
AC Q8WZ58;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Achaete-scute associated protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21334642; PubMed=11440538;
RA Westerman B.A., Poutsma A., Looijenga L.H., Wouters D., van Wijk I.J.,
RA Oudejans C.B.;
RT "The Human Achaete Scute Homolog 2 gene contains two promoters,
RT generating overlapping transcripts and encoding two proteins with
RT different nuclear localization.";
RL Placenta 22:511-518(2001).
DR EMBL; AF442769; AAL35361.1; -.
SQ SEQUENCE 368 AA; 39090 MW; 2D4E7CDF80CC8A39 CRC64;

Query Match 81.4%; Score 35; DB 4; Length 368;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRVRPH 7

```

```

DB 346 GRVRPH 351

RESULT 11
Q8S6L5 PRELIMINARY; PRT; 404 AA.
AC Q8S6L5;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical 43.8 kDa protein.
GN OSJNBA0019N10.11.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIPPONBARE;
RA McCombie W.R., de la Bastide M., Spiegel L., Preston R., Kirchoff K.,
RA Kuit K., Nascimento L., Zutavern T., Ballija V., Bell M., Baker J.,
RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,
RA O'Shaughnessy A., Palmer L., Dedhia N.;
RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone
RT OSJNBA0019N10, from chromosome 10, complete sequence.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC092748; AAM01073.1; -.
KW Hypothetical protein.
SQ SEQUENCE 404 AA; 43843 MW; C5D1E355ECBE6599 CRC64;

Query Match 81.4%; Score 35; DB 10; Length 404;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GRVRPH 7
DB 86 GRVRPH 91

RESULT 12
Q9E942 PRELIMINARY; PRT; 437 AA.
AC Q9E942;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE ORF437 protein.
GN LOF2.
OS Gallid herpesvirus 3.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=35250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HPRS24;
RA Izumiya Y., Jang H., Ono M., Mikami T.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HPRS24;
RA Izumiya Y., Jang H., Ono M., Mikami T.;
RT "A Complete Genomic DNA Sequence of Marek's Disease Virus Type 2,
RT Strain HPRS24.";
RL Curr. Top. Microbiol. Immunol. 0:0-0(2000).
DR EMBL; AB049735; BAB16507.1; -. 758DD400F5CA6AFB CRC64;
SQ SEQUENCE 437 AA; 48879 MW; 758DD400F5CA6AFB CRC64;

Query Match 81.4%; Score 35; DB 12; Length 437;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 2 GVRPH 7
 Db 226 GVRPH 231

RESULT 13

Q9FTM3 PRELIMINARY; PRT; 438 AA.
 AC Q9FTM3
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DE P0005A05.26 protein (P0482C06.9 protein).
 GN P0005A05.26 OR P0482C06.9.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoidae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN-CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 clone:P0005A05.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RC SEQUENCE FROM N.A.
 RC STRAIN-CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 clone:P0482C06.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP002863; BAB16222.1;
 DR EMBL; AP002845; BAB78613.1;
 DR InterPro; IPR004263; Exostosin.
 DR Pfam; PF03016; Exostosin; 1.
 SQ SEQUENCE 438 AA; 49435 MW; 256DE806D01D41E7 CRC64;

Query Match 81.4%; Score 35; DB 10; Length 438;
 Best Local Similarity 100.0%; Pred. No. 81;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGRVRP 6
 Db 283 HGRVRP 288

RESULT 14

O69864 PRELIMINARY; PRT; 442 AA.
 AC O69864
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DE Hypothetical protein SCO6034.
 GN SCO6034 OR SC1C3.22
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN-A3(2) / M145;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares K., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;

RT *Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 DR EMBL; AL023702; CAAL9246.1;
 KW Hypothetical protein.
 SQ SEQUENCE 442 AA; 48524 MW; F3821B9CB50299EC CRC64;

Query Match 81.4%; Score 35; DB 16; Length 442;
 Best Local Similarity 100.0%; Pred. No. 82;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGRVRP 6
 Db 349 HGRVRP 354

RESULT 15

O9DE31 PRELIMINARY; PRT; 534 AA.
 AC O9DE31
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DE TGF-beta type II receptor (Fragment).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX MEDLINE=21125228; PubMed=11232255;
 RA Dhanasekaran S.M., Vempati U.D., Kondaliah P.;
 RT "Isolation and characterization of a transforming growth factor-beta
 Type II receptor cDNA from Xenopus laevis.";
 RL Gene 263:171-178(2001).
 DR EMBL; AF213685; AAG40577.1;
 DR InterPro; IPR000333; Actn_receptorII.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00653; ACTIVIN2R.
 DR PRODOM; PD000001; Euk_pkinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00219; TyKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW Receptor.
 FT NON_TER
 SQ SEQUENCE 534 AA; 60509 MW; B4F48C11C30C21EF CRC64;

Query Match 81.4%; Score 35; DB 13; Length 534;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGRVRP 6
 Db 472 HGRVRP 477

Search completed: July 16, 2003, 13:15:11
 Job time : 20.9512 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:04:13 ; Search time 23.9024 Seconds
(without alignments)
39.023 Million cell updates/sec

Title: US-09-910-582b-4
Perfect score: 29
Sequence: 1 VLVVTS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
1	29	100.0	7	22 AAB30897 Peptide which sele
2	29	100.0	7	22 AAB50799 Heart homing pepti
3	28	96.6	277	22 AAB96795 Putative P. abyssi
4	27	93.1	84	22 AAU39832 Propionibacterium
5	27	93.1	126	23 ABP38921 Staphylococcus epi
6	27	93.1	183	22 ABG62077 Drosophila melanog
7	27	93.1	545	22 ABG28901 Novel human diagn
8	27	93.1	586	21 ABG39789 Arabidopsis thalia
9	26	89.7	64	23 ABP07206 Human ORFX protein
10	26	89.7	83	22 AAG91396 C glutamicum prote

11	26	89.7	155	22 AAG72985 Olfactory receptor
12	26	89.7	177	22 AAB62330 Drosophila melanog
13	26	89.7	188	22 AAG91591 C glutamicum prote
14	26	89.7	244	21 AAG58124 Arabidopsis thalia
15	26	89.7	251	22 ABG27794 Novel human diagn
16	26	89.7	256	21 AAG58123 Arabidopsis thalia
17	26	89.7	291	21 AAG09408 Arabidopsis thalia
18	26	89.7	310	21 AAB26749 Human beta-1,3 gal
19	26	89.7	310	21 AAY94641 Human beta3Gal-T5
20	26	89.7	344	21 AAG09407 Arabidopsis thalia
21	26	89.7	361	21 AAG09406 Arabidopsis thalia
22	26	89.7	710	22 AAB30814 Amino acid sequenc
23	26	89.7	730	21 AAG39220 Arabidopsis thalia
24	26	89.7	736	22 AAB20185 Human multifunctio
25	26	89.7	743	21 AAG30752 Arabidopsis thalia
26	26	89.7	785	21 AAG39013 Arabidopsis thalia
27	26	89.7	797	21 AAG39219 Arabidopsis thalia
28	26	89.7	799	22 AAB94351 Human protein sequ
29	26	89.7	810	21 AAG30751 Arabidopsis thalia
30	26	89.7	818	21 AAG39218 Arabidopsis thalia
31	26	89.7	831	21 AAG30750 Arabidopsis thalia
32	26	89.7	852	21 AAG39012 Arabidopsis thalia
33	26	89.7	873	21 AAG39011 Arabidopsis thalia
34	26	89.7	874	22 AAB93264 Human protein sequ
35	26	89.7	879	22 AAB58217 Drosophila melanog
36	26	89.7	904	22 AAO13497 Human polypeptide
37	25	86.2	14	22 AAM98874 Human peptide #214
38	25	86.2	88	23 AAB31493 Human ORF466 prote
39	25	86.2	104	18 AAW27862 Amino acid sequenc
40	25	86.2	108	23 ABP40411 Staphylococcus epi
41	25	86.2	139	19 AAW57446 Anti-CD4 antibody
42	25	86.2	148	22 AAU54893 Propionibacterium
43	25	86.2	169	22 AAU53901 Propionibacterium
44	25	86.2	228	22 AAB19994 Rat camello 3 (Rcm
45	25	86.2	22	22 AAU37146 Staphylococcus aur

ALIGNMENTS

RESULT 1

AAB30897 ID AAB30897 standard; peptide; 7 AA.

XX AC AAB30897;

XX DT 02-APR-2001 (first entry)

XX DE Peptide which selectively binds to normal cardiac endothelium.

XX DE Cardiac endothelium; angiogenic factor; vascular endothelium;

XX KW peripheral vascular disease; cardiovascular disease; angiogenesis;

XX KW cardiac neovascularisation.

XX OS Unidentified.

XX OS WO200075329-A1.

XX PD 14-DEC-2000.

XX PF 31-MAY-2000; 2000WO-US14988.

XX PF 07-JUN-1999; 99US-0327045.

XX PA (EDMA-) EDWARDS LIFESCIENCES CORP.

XX PA Putative P. abyssi

XX PI Propionibacterium

XX PI Staphylococcus epi

XX PI Drosophila melanog

XX PI Novel human diagn

XX PI Arabidopsis thalia

XX PT Human ORFX protein

XX PT C glutamicum prote

PT New chimeric molecules having an angiogenic factor linked to a

PT targeting molecule that binds to a vascular endothelium, useful for

Wed Jul 16 13:41:36 2003

PT increasing cardiac neovascularisation, or treating peripheral vascular
PT and cardiovascular diseases -

XX Disclosure; Page 27; 67pp; English.

XX AAB30895-99 represent targeting molecules, which are used to produce the
CC chimeric molecules of the invention. AAB30895-98 selectively bind to
CC normal cardiac endothelium. The specification describes a chimeric
CC molecule comprising an angiogenic factor linked to a targeting molecule
CC that specifically binds to a vascular endothelium. The chimeric
CC molecules are useful for treatment of peripheral vascular or
CC cardiovascular diseases. Specifically, they are useful for inducing or
CC inhibiting angiogenesis, for increasing cardiac neovascularisation in
CC ischemic tissue in the peripheral vascular system.

XX Sequence 7 AA;

Query Match 100.0%; Score 29; DB 22; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVLVTSS 7
DB 1 VVLVTSS 7

RESULT 2

AAB50799
ID AAB50799 standard; peptide; 7 AA.

XX AAB50799;

XX 21-MAR-2001 (first entry)

XX Heart homing peptide SEQ ID NO: 4.

XX Heart homing peptide; cardiovascular disease; ischaemic disease;

XX gene therapy.

XX Synthetic.

XX WO200075174-A1.

XX 14-DEC-2000.

XX 31-MAY-2000; 2000WO-US15088.

XX 07-JUN-1999; 99US-0326718.

XX (BURN-) BURNHAM INST.

XX Ruoslahti E, Mackenna DA;

XX WPI; 2001-071059/08.

XX Novel heart homing peptide that selectively homes to normal ischaemic
XX and cardiac tissue useful for targeting ischaemic tissues for treating
XX ischaemic and cardiovascular diseases such as atherosclerosis and
XX restenosis

XX Claim 2; Page 55; 70pp; English.

XX The present invention provides a number of heart homing peptides which
XX selectively home to cardiac tissue. These can be used in the treatment of
XX cardiovascular and ischaemic diseases, such as atherosclerosis,
XX thrombosis, restenosis, vasculitis, atherosclerotic aneurysms, myocardial
XX hypertrophy, congenital heart diseases, ischaemic heart disease and
XX anginas, acquired valvular/endocardial diseases, primary myocardial
XX diseases, cardiac tumours and arrhythmias.

XX Sequence 7 AA;

Query Match 100.0%; Score 29; DB 22; Length 7;

PT increasing cardiac neovascularisation, or treating peripheral vascular
PT and cardiovascular diseases -

XX Disclosure; Page 27; 67pp; English.

XX AAB30895-99 represent targeting molecules, which are used to produce the
CC chimeric molecules of the invention. AAB30895-98 selectively bind to
CC normal cardiac endothelium. The specification describes a chimeric
CC molecule comprising an angiogenic factor linked to a targeting molecule
CC that specifically binds to a vascular endothelium. The chimeric
CC molecules are useful for treatment of peripheral vascular or
CC cardiovascular diseases. Specifically, they are useful for inducing or
CC inhibiting angiogenesis, for increasing cardiac neovascularisation in
CC ischemic tissue in the peripheral vascular system.

XX Sequence 7 AA;

Query Match 100.0%; Score 29; DB 22; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVLVTSS 7
DB 1 VVLVTSS 7

RESULT 3

AAB96795
ID AAB96795 standard; Protein; 277 AA.

XX AAB96795;

XX 29-OCT-2001 (first entry)

XX Putative P. abyssal permease #29.

XX Hyperthermophilic archaeon; hyperthermophilic protein.

XX Pyrococcus abyssal.

XX FR2792651-A1.

XX 27-OCT-2000.

XX 21-APR-1999; 99FR-0005034.

XX 21-APR-1999; 99FR-0005034.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX (IFRE-) IFREMER INST FR RECH EXPL MER.

XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
XX Querellou J, Weissbach J, Saurin W, Heilig R;

XX WPI; 2001-126236/14.

XX New nucleotide sequences isolated from Pyrococcus abyssal encode
XX proteins useful in industry -

XX Claim 7; Pages 1568-1569; 1657pp; French.

XX The present invention relates to the genomic sequence of Pyrococcus
XX abyssal (see AAF86431 and AAH41223-7) and P. abyssal proteins. P. abyssal is
XX a hyperthermophilic archaeon, which is isolated from deep-sea
XX hydrothermal vents. The present sequence is one such P. abyssal protein.
XX The proteins of the present invention have various potential industrial
XX uses, since the proteins are stable at very high temperatures, some up to
XX 110 degrees centigrade.

XX Note: This patent is in the same patent family as WO200065062, which
XX contains additional sequences as shown in AAB99132-AA999143,
XX AAH75903-AAH75920 and AAG66436.

XX Sequence 277 AA;

Query Match 96.6%; Score 28; DB 22; Length 277;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVLVTSS 7
DB 155 VVLVTSS 161

RESULT 4

AAU39832
ID AAU39832 standard; Protein; 84 AA.

XX AAU39832;

XX 13-FEB-2002 (first entry)

XX

DE Propionibacterium acnes immunogenic protein #728.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 PN
 PN W0200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI; 2001-616774/71.
 DR N-PSDB; AAS59508.
 XX
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 PS Example 1; SEQ ID NO 1027; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 84 AA;
 Query Match 93.1%; Score 27; DB 22; Length 84;
 Best Local Similarity 85.7%; Pred. No. 59;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VVLVTSS 7
 DB 61 VVMVTSS 67
 RESULT 5
 ID ABP38921 standard; Protein; 126 AA.
 XX
 AC ABP38921;
 XX
 DT 24-JUL-2002 (first entry)
 XX

DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3766.
 XX
 KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
 KW antibacterial; gene therapy.
 XX
 OS Staphylococcus epidermidis.
 PN
 PN US6380370-B1.
 XX
 PD 30-APR-2002.
 XX
 PF 13-AUG-1998; 98US-0134001.
 XX
 PR 14-AUG-1997; 97US-055779P.
 PR 08-NOV-1997; 97US-064964P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Doucette-Stamm LA, Bush D;
 XX
 DR WPI; 2002-381255/41.
 DR N-PSDB; ABN91466.
 XX
 PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis
 PT polypeptide, useful for diagnosing and treating bacterial infections -
 XX
 PS Disclosure; SEQ ID 3766; 267pp; English.
 XX
 CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
 CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
 CC antibacterial activity and can be used in gene therapy. The sequences
 CC can also be used in the diagnosis and treatment of bacterial infections,
 CC particularly S. epidermidis infections. The sequences can be used to
 CC screen for compounds able to interfere with the S. epidermidis life
 CC cycle or inhibit S. epidermidis infection.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC USPTO web site.
 XX
 SQ Sequence 126 AA;
 Query Match 93.1%; Score 27; DB 23; Length 126;
 Best Local Similarity 71.4%; Pred. No. 96;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VVLVTSS 7
 DB 115 VVLVTSS 121
 RESULT 6
 ID ABP62077 standard; Protein; 183 AA.
 XX
 AC ABP62077;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 13023.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN W0200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX

PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
PA Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
XX N-PSDB; ABL06180.
DR New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX Disclosure; SEQ ID NO 13023; 21pp + Sequence Listing; English.
PS The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (AB95737-AB972072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 183 AA;
Query Match 93.1%; Score 27; DB 22; Length 183;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLVTS 7
DB 35 IILVTS 41
RESULT 7
ABG28901
ID ABG28901 standard; Protein; 545 AA.
XX AC ABG28901;
XX 18-FEB-2002 (first entry)
DE Novel human diagnostic protein #2892.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
OS WO200175067-A2.
PN 11-OCT-2001.
PD 30-MAR-2001; 2001WO-US08631.
PF 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
PI WPI; 2001-639362/73.
XX N-PSDB; AAS93088.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

XX Claim 20; SEQ ID NO 59260; 103pp; English.
PS The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 545 AA;
Query Match 93.1%; Score 27; DB 22; Length 545;
Best Local Similarity 85.7%; Pred. No. 5.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLVTS 7
DB 193 VLVTS 199
RESULT 8
AAG39789
ID AAG39789 standard; Protein; 588 AA.
XX AC AAG39789;
XX 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 49284.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS EP1033405-A2.
PN 06-SEP-2000.
PD 25-FEB-2000; 2000EP-0301439.
PF 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 05-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142377.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 18-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 28-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

RESULT 10	
AAG91396	
ID	AAG91396 standard; Protein; 83 AA.
XX	
AC	AAG91396;
XX	
XX	
DT	26-SEP-2001 (first entry)
XX	
DE	C glutamicum protein fragment SEQ ID NO: 5150.
XX	
XX	Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW	organic acid synthesis.
KW	
XX	
OS	Corynebacterium glutamicum.
XX	
PN	EP1108790-A2.
XX	
PD	20-JUN-2001.
XX	
PF	18-DEC-2000; 2000EP-0127588.
XX	
PR	16-DEC-1999; 99JP-0377484.
PR	07-APR-2000; 2000JP-0159162.
PR	03-AUG-2000; 2000JP-0280988.
XX	
PA	(KYOW) KYOWA HAKKO KOGYO KK.
XX	
PI	Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI	Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX	
DR	WPI: 2001-376931/40.
DR	N-PSDB; AAH66615.
XX	
XX	Mutational polynucleotides derived from Coryneform bacteria, for identifying
PT	mutation point of a gene, measuring expression of a gene, analysing
PT	expression profile or pattern of a gene and identifying homologous gene
PT	

24-JUN-2002 (first entry)

Human ORFX protein sequence SEQ ID NO:14394.

Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis; hyperproliferative disorder; psoriasis; benign tumour; haemorrhage; degenerative disorder; osteoarthritis; neurodegenerative disorder; tosus; cardiovascular disease; diabetes mellitus; systemic lupus erythematosus; hypertension; hypothyroidism; cholesterol ester storage disease; immune deficiency; immune disorder; infectious disease; autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis; myasthenia gravis.

Homo sapiens.

WO200192523-A2.

06-DEC-2001.

29-MAY-2001; 2001WO-US10836.

30-MAY-2000; 2000US-206132P.

29-AUG-2000; 2000US-228716P.

(CURA) CURAGEN CORP.

Shinkets RA, Leach MD;

WPI; 2002-106308/14.

N-PSDB; ABN22958.

Novel human polypeptides and polynucleotides useful for diagnosing, preventing and treating cardiovascular disease, neurodegenerative, hyperproliferative disorders and autoimmune disorders

Disclosure; SEQ ID 14394; 1037pp; English.

XX PS Claim 17; SEQ ID NO: 5150; 246pp + Sequence Listing; English.

CC The present invention provides a number of nucleotide and protein sequences from the Coryneform bacterium *Corynebacterium glutamicum*. These are useful for identifying the mutation point of a gene derived from a mutant of coryneform bacterium, measuring expression amount and analysing the expression profile or expression pattern of a gene derived from Coryneform bacterium, and identifying a homologue of a gene derived from coryneform bacterium. Coryneform bacteria are useful for producing amino acids, nucleic acids, vitamins, saccharides and organic acids, particularly L-lysine. The present sequence is a protein described in the exemplification of the invention.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the European Patent Office.

XX SQ Sequence 83 AA;

Query Match 89.7%; Score 26; DB 22; Length 83;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
Db 43 VLVLTSA 49
|||||

RESULT 11

AAG72985

ID AAG72985 standard; Protein; 155 AA.

XX AC AAG72985;

XX DT 30-JUL-2001 (first entry)

XX DE Olfactory receptor-like polypeptide, SEQ ID NO: 2667.

XX KW Human; olfactory receptor; OR; primary scent determination;
secondary scent determination; polypeptide library; odour receptor;
scent profile; scent fingerprint; scent representation.

XX OS Unidentified.

XX PN WO200127158-A2.

XX PD 19-APR-2001.

XX PF 06-OCT-2000; 2000WO-US27582.

XX PR 08-OCT-1999; 99US-0158615.

XX PR 24-FEB-2000; 2000US-0184809.

XX PA (DIGI-) DIGISCENTS.

XX PA (YEDA) YEDA RES & DEV CO LTD.

XX PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;

XX PF WI; 2001-290713/30.

XX DR New polynucleotides which encode polypeptides involved in olfactory sensation for identifying olfactory agonists and antagonists -

XX PS Example 6; Page 1803-1804; 1857pp; English.

XX CC The present sequence is an olfactory receptor polypeptide which was used as a query sequence in a database search of olfactory receptor (OR)-like sequences. The invention relates to isolated polynucleotides encoding polypeptides involved in olfactory sensation. The polynucleotides can be used in screening for olfactory agonists and antagonists. The methods allow for the determination of primary scents and the identification of the odour receptors used to detect these primary scents. The methods also enable determination of secondary scents

CC and the identification of combinations of odour receptors that are involved in detecting such secondary scents. This enables the construction of a scent representation (also called a scent fingerprint or scent profile), which may be used to re-create and edit scents. Libraries of olfactory receptors are useful for determining the interaction pattern of a composition with the receptors, and can be used for determining differences in the olfactory faculties of different individuals.

XX SQ Sequence 155 AA;

Query Match 89.7%; Score 26; DB 22; Length 155;
Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
Db 76 VLLTSS 82
|||||

RESULT 12

ABB62330

ID ABB62330 standard; Protein; 177 AA.

XX AC ABB62330;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 13782.

XX KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL06433.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

XX PS Disclosure; SEQ ID NO 13782; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).

XX CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 177 AA;

Query Match 89.7%; Score 26; DB 22; Length 177;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;

Wed Jul 16 13:41:36 2003

us-09-910-582b-4.rag

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VVLVTSS 7
I:|||||
Db 34 VLLVTSS 40

RESULT 13

AAG91591
ID AAG91591 standard; Protein: 188 AA.

XX
AC AAG91591;

DT 26-SEP-2001 (first entry)

XX C glutamicum protein fragment SEQ ID NO: 5345.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.

XX Corynebacterium glutamicum.

OS
XX EP1108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-0127688.

XX 16-DEC-1999; 99JP-0377484.

PR 07-APR-2000; 2000JP-0159162.

PR 03-AUG-2000; 2000JP-0280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI: 2001-376931/40.

DR N-PSDB; AAH66810.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene

XX Claim 17; SEQ ID NO: 5345; 246pp + Sequence Listing: English.

XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.

XX Sequence 188 AA;

Query Match 89.7%; Score 26; DB 22; Length 188;

Best Local Similarity 57.1%; Pred. No. 2.6e+02;

Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVLVTSS 7

Db 16 ILLVTSS 22

RESULT 14

AAG58124
ID AAG58124 standard; Protein: 244 AA.

XX
AC AAG58124;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 74992.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136382.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 89.7%; Score 26; DB 21; Length 244;

Best Local Similarity 71.4%; Pred. No. 3.5e+02;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVLVTSS 7
1::1111
Db 3 VIIVTSS 9

RESULT 15

ABG27794

ID ABG27794 standard; Protein; 251 AA.

XX ABG27794;

AC ABG27794;

XX 18-FEB-2002 (first entry)

XX

Wed Jul 16 13:41:36 2003

DE Novel human diagnostic protein #27785.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 OS WO200175067-A2.
 PN 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 XX 23-AUG-2000; 2000US-0649167.
 PR (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 XX N-PSDB; AAS91981.
 DR New isolated polynucleotide and encoded polypeptides, useful in
 XX diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 PT
 XX Claim 20; SEQ ID No 58153; 103pp; English.
 PS
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 251 AA;
 Query Match 89.7%; Score 26; DB 22; Length 251;
 Best Local Similarity 85.7%; Pred. No. 3.7e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VVLTSS 7
 III:III
 DB 13 VVLTSS 19
 Search completed: July 16, 2003, 13:12:39
 Job time : 24.9024 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:10:13 ; Search time 8.19512 Seconds
(without alignments)
25.132 Million cell updates/sec

Title: US-09-910-582b-4
Perfect score: 29
Sequence: 1 VVLVTSS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2.6/ptodata/1/1aa/5A-COMB.pep:*
2: /cgn2.6/ptodata/1/1aa/5B-COMB.pep:*
3: /cgn2.6/ptodata/1/1aa/6A-COMB.pep:*
4: /cgn2.6/ptodata/1/1aa/6B-COMB.pep:*
5: /cgn2.6/ptodata/1/1aa/6C-COMB.pep:*
6: /cgn2.6/ptodata/1/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	100.0	7	US-09-326-718-4	Sequence 4, Appli
2	27	93.1	126	US-09-134-001C-3766	Sequence 3766, Ap
3	25	86.2	108	US-09-134-001C-5256	Sequence 5256, Ap
4	25	86.2	247	US-08-961-083-170	Sequence 170, Appl
5	25	86.2	554	US-09-321-276-4	Sequence 4, Appli
6	25	86.2	554	US-08-916-481-2	Sequence 2, Appli
7	25	86.2	563	US-08-916-481-3	Sequence 3, Appli
8	24	82.8	165	US-08-858-207A-520	Sequence 520, App
9	24	82.8	199	PCT-US93-10418-4	Sequence 4, Appli
10	24	82.8	326	US-08-306-511A-10	Sequence 10, Appl
11	24	82.8	326	US-08-893-274-10	Sequence 10, Appl
12	24	82.8	326	US-08-581-918A-10	Sequence 10, Appl
13	24	82.8	326	US-08-346-147B-10	Sequence 10, Appl
14	24	82.8	326	US-08-822-936-10	Sequence 10, Appl
15	24	82.8	326	US-08-497-214D-10	Sequence 10, Appl
16	24	82.8	326	PCT-US95-04636-10	Sequence 10, Appl
17	24	82.8	331	US-08-258-188-5	Sequence 5, Appli
18	24	82.8	331	US-08-526-813-5	Sequence 5, Appli
19	24	82.8	331	PCT-US95-08554-5	Sequence 5, Appli
20	24	82.8	450	US-08-788-800-12	Sequence 12, Appl
21	24	82.8	469	US-07-934-373C-23	Sequence 23, Appl
22	24	82.8	469	US-08-437-642B-23	Sequence 23, Appl
23	24	82.8	469	US-08-146-206C-23	Sequence 23, Appl
24	24	82.8	532	US-08-140-729A-3	Sequence 3, Appli
25	24	82.8	532	US-08-546-666-3	Sequence 3, Appli
26	24	82.8	532	US-08-916-745-3	Sequence 3, Appli
27	24	82.8	532	US-09-042-929-3	Sequence 3, Appli

28 24 82.8 532 2 US-08-546-661-3 Sequence 3, Appli
29 24 82.8 532 2 US-09-042-960-3 Sequence 3, Appli
30 24 82.8 532 3 US-09-198-650-3 Sequence 3, Appli
31 24 82.8 532 3 US-09-042-913-3 Sequence 3, Appli
32 24 82.8 532 3 US-09-042-937-3 Sequence 3, Appli
33 24 82.8 532 4 US-09-042-709A-3 Sequence 23, Appli
34 24 82.8 552 5 PCT-US93-07832-23 Sequence 23, Appli
35 24 82.8 820 2 US-08-380-182-23 Sequence 23, Appli
36 23 79.3 27 2 US-08-103-170-15 Sequence 15, Appli
37 23 79.3 36 2 US-08-411-098-15 Sequence 15, Appli
38 23 79.3 219 4 US-09-134-001C-4581 Sequence 4581, Ap
39 23 79.3 344 4 US-09-389-331-72 Sequence 72, Appli
40 23 79.3 513 2 US-08-357-533A-11 Sequence 11, Appli
41 23 79.3 513 2 US-08-459-009-11 Sequence 11, Appli
42 23 79.3 513 3 US-08-459-951-11 Sequence 11, Appli
43 23 79.3 567 3 US-08-813-150-4 Sequence 4, Appli
44 23 79.3 605 4 US-09-134-001C-4425 Sequence 4425, Ap
45 23 79.3 625 1 US-08-365-981-13 Sequence 13, Appli

ALIGNMENTS

RESULT 1
US-09-326-718-4
; Sequence 4, Application US/09326718
; Patent No. 6303573
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Mackenna, Deidre A.
; TITLE OF INVENTION: Heart Homing Peptides and Methods of
; TITLE OF INVENTION: Using Same
; FILE REFERENCE: P-LJ 3512
; CURRENT APPLICATION NUMBER: US/09/326,718
; CURRENT FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-09-326-718-4

Query Match 100.0%; Score 29; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VVLVTSS 7
DB 1 VVLVTSS 7
RESULT 2
US-09-134-001C-3766
; Sequence 3766, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: CTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3766
; LENGTH: 126
; TYPE: PRT

QY 1[~] VLVLTSS 7
99 VLVLTSS 105
nb

RESULT 6
US-08-916-481-2
; Sequence 2, Application US/08916481
; Patent No. 6270991
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; TITLE OF INVENTION: NOVEL histidine kinase
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,481
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd Q
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: GM10022-1
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 554 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-916-481-2

Query Match 86.2%; Score 25; DB 4; Length 554;
Best Local Similarity 85.7%; Pred. No. 7.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLVLTSS 7
Db 99 VLVLTSS 105

RESULT 7
US-08-916-481-3
; Sequence 3, Application US/08916481
; Patent No. 6270991
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; TITLE OF INVENTION: NOVEL histidine kinase
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,481
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd Q
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: GM10022-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 563 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-916-481-3

Query Match 86.2%; Score 25; DB 4; Length 563;
Best Local Similarity 85.7%; Pred. No. 7.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLVLTSS 7
Db 108 VLVLTSS 114

RESULT 8
US-08-858-207A-520
; Sequence 520, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 520:
; SEQUENCE CHARACTERISTICS:

us-09-910-582b-4.ra1

Wed Jul 16 13:41:36 2003

us-09-910-582b-4.ra1

```

;
; LENGTH: 165 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6348328e
;
US-08-858-207A-520
;
Query Match 82.8%; Score 24; DB 4; Length 165;
Best Local Similarity 71.4%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLVTS 7
Db 85 VLVTS 91

RESULT 9
PCT-US93-10418-4
; Sequence 4, Application PC/TUS9310418
; GENERAL INFORMATION:
; APPLICANT: Ziegler, Steven F.
; APPLICANT: Hjerrild, Kathryn A.
; TITLE OF INVENTION: Activation Antigen CD69
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
;
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10418
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2610-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 199 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
PCT-US93-10418-4
;
Query Match 82.8%; Score 24; DB 5; Length 199;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVTS 6
Db 48 VLVTS 53

RESULT 10
US-08-306-511A-10
; Sequence 10, Application US/08306511A
; Patent No. 5962316
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Demetrick, Douglas J.
; APPLICANT: Serrano, Manuel
; APPLICANT: Hannon, Gregory J.
;
PCT-US93-10418-4
;
TITLE OF INVENTION: Cell-Cycle Regulatory Proteins, and Uses
; TITLE OF INVENTION: Related Thereto
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
;
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306,511A
; FILING DATE: 14-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSI-001CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 326 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: N-terminal
;
US-08-306-511A-10
;
Query Match 82.8%; Score 24; DB 2; Length 326;
Best Local Similarity 83.3%; Pred. No. 7.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLVTS 7
Db 151 ILVTS 156

RESULT 11
US-08-893-274-10
; Sequence 10, Application US/08893274
; Patent No. 5968821
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Demetrick, Douglas J.
; APPLICANT: Serrano, Manuel
; APPLICANT: Hannon, Gregory J.
; TITLE OF INVENTION: Cell-Cycle Regulatory Proteins,
; TITLE OF INVENTION: and Uses Related Thereto
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
;
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/893,274
; FILING DATE: 15-JULY-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,511
```

; FILING DATE: 14-SEPTEMBER-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/248,812
; FILING DATE: 25-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/227,371
; FILING DATE: 14-APRIL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/154,915
; FILING DATE: 18-NOVEMBER-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/991,997
; FILING DATE: 17-DECEMBER-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCTOBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-071.09
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 326 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: N-terminal
; US-08-893-274-10

Query Match 82.8%; Score 24; DB 2; Length 326;
Best Local Similarity 83.3%; Pred. No. 7.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 VLVTS 7
:|||||
Db 151 ILVTS 156

RESULT 12
US-08-581-918A-10
; Sequence 10, Application US/08581918A
; Patent No. 6043030
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Demetrick, Douglas J.
; APPLICANT: Serrano, Manuel
; APPLICANT: Hannon, Gregory J.
; TITLE OF INVENTION: Cell-Cycle Regulatory Proteins, and Uses
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Elliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPad
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/581,918A
; FILING DATE: 02-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/497,214
; FILING DATE: 30-JUN-1995
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/346,147
; FILING DATE: 29-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,511
; FILING DATE: 14-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/248,812
; FILING DATE: 25-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/227,371
; FILING DATE: 14-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/154,915
; FILING DATE: 18-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/991,997
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-071.06
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1299
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 326 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: N-terminal
; US-08-581-918A-10

Query Match 82.8%; Score 24; DB 3; Length 326;
Best Local Similarity 83.3%; Pred. No. 7.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 VLVTS 7
:|||||
Db 151 ILVTS 156

RESULT 13
US-08-346-147B-10
; Sequence 10, Application US/08346147B
; Patent No. 6211334
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Demetrick, Douglas J.
; APPLICANT: Serrano, Manuel
; APPLICANT: Hannon, Gregory J.
; TITLE OF INVENTION: Cell-Cycle Regulatory Proteins, and Uses
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Elliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPad
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/346,147B
; FILING DATE: 29-NOV-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,511
; FILING DATE: 14-SEP-1994

Wed Jul 16 13:41:36 2003

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 326 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: N-terminal
; US-08-822-936-10
;
; Query Match      82.8%; Score 24; DB 4; Length 326;
; Best Local Similarity 83.3%; Pred. No. 7.4e+02;
; Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
;
QY      2 VLVTS 7
Db      151 ILVTS 156

RESULT 15
US-08-497-214D-10
; Sequence 10, Application US/08497214D
; Patent No. 6331390
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Demetrick, Douglas J.
; APPLICANT: Serrano, Manuel
; APPLICANT: Hannon, Gregory J.
; TITLE OF INVENTION: Cell-Cycle Regulatory Proteins, and Uses
; TITLE OF INVENTION: Related Thereto
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Elliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordpad
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/497,214D
; FILING DATE: 30-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/346,147
; FILING DATE: 29-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,511
; FILING DATE: 14-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/248,812
; FILING DATE: 25-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/227,371
; FILING DATE: 14-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/154,915
; FILING DATE: 18-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/991,997
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-071.04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1299
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 326 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: N-terminal
; US-08-346-147B-10
;
; Query Match      82.8%; Score 24; DB 4; Length 326;
; Best Local Similarity 83.3%; Pred. No. 7.4e+02;
; Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
;
QY      2 VLVTS 7
Db      151 ILVTS 156

RESULT 14
US-08-822-936-10
; Sequence 10, Application US/08822936
; Patent No. 6242575
; GENERAL INFORMATION:
; APPLICANT: Massague, Joan
; APPLICANT: Roberts, James M.
; APPLICANT: Koff, Andrew
; APPLICANT: Polyak, Kornelia
; TITLE OF INVENTION: Isolated p27 Protein, Nucleic Acid
; TITLE OF INVENTION: Molecules Encoding Same, Methods of Identifying Agents Acting
; TITLE OF INVENTION: and Uses of Said Agents
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Elliot, LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,936
; FILING DATE: 21-FEBRUARY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-079.05
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 10:

```


; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: N-terminal
US-08-497-214D-10

Query Match 82.8%; Score 24; DB 4; Length 326;
Best Local Similarity 83.3%; Pred. NO. 7.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLVTS 7
Db 151 ILVTS 156

Search completed: July 16, 2003, 13:17:07
Job time : 9.19512 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:15:24 ; Search time 13.3171 Seconds
(without alignments)
62.425 Million cell updates/sec

Title: US-09-910-582B-4
Perfect score: 29
Sequence: 1 VVLVTSS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	29	100.0	7	9	US-09-910-582B-4	Sequence 4, Appli
2	29	100.0	7	10	US-09-782-650-3	Sequence 3, Appli
3	29	100.0	498	9	US-10-128-714-3581	Sequence 3581, Ap
4	29	100.0	522	9	US-10-128-714-8581	Sequence 8581, Ap
5	26	89.7	83	9	US-09-738-626-5150	Sequence 5150, Ap
6	26	89.7	188	9	US-09-738-626-5345	Sequence 5345, Ap
7	26	89.7	310	9	US-10-344-440-1	Sequence 1, Appli
8	26	89.7	736	9	US-10-060-230-24	Sequence 24, Appl
9	26	89.7	874	9	US-10-205-823-232	Sequence 232, App
10	25	86.2	243	10	US-09-815-242-12739	Sequence 12739, A
11	25	86.2	247	10	US-09-765-272-170	Sequence 170, App
12	25	86.2	272	9	US-10-156-761-10426	Sequence 10426, A
13	25	86.2	295	9	US-10-127-032-115	Sequence 115, App
14	25	86.2	299	10	US-09-861-451A-36	Sequence 36, Appl
15	25	86.2	343	9	US-10-156-761-9520	Sequence 9520, Ap
16	25	86.2	350	9	US-10-156-761-12017	Sequence 12017, A
17	25	86.2	410	9	US-10-050-704-141	Sequence 141, App
18	25	86.2	457	10	US-09-815-242-11616	Sequence 11616, A
19	25	86.2	553	10	US-09-815-242-5778	Sequence 5778, Ap

20	25	86.2	554	9	US-09-878-672-2	Sequence 2, Appli
21	25	86.2	554	10	US-09-800-396-4	Sequence 4, Appli
22	25	86.2	563	9	US-09-878-672-3	Sequence 3, Appli
23	25	86.2	563	9	US-09-769-787-27	Sequence 27, Appl
24	25	86.2	563	9	US-09-769-787-28	Sequence 28, Appl
25	25	86.2	574	10	US-09-815-242-12045	Sequence 12045, A
26	25	86.2	604	10	US-09-815-242-10022	Sequence 10022, A
27	25	86.2	906	9	US-10-060-230-21	Sequence 21, Appl
28	25	86.2	906	9	US-10-060-230-22	Sequence 22, Appl
29	25	86.2	1014	9	US-10-054-044A-4	Sequence 4, Appli
30	25	86.2	1014	9	US-10-120-604-4	Sequence 4, Appli
31	25	86.2	1014	9	US-10-120-604-139	Sequence 139, App
32	25	86.2	1014	9	US-10-225-567A-428	Sequence 428, App
33	24	82.8	33	10	US-09-864-761-34372	Sequence 34372, A
34	24	82.8	53	10	US-09-864-761-38747	Sequence 38747, A
35	24	82.8	138	9	US-10-156-761-9023	Sequence 9023, Ap
36	24	82.8	167	9	US-10-080-170-214	Sequence 214, App
37	24	82.8	205	9	US-09-738-626-5902	Sequence 5902, Ap
38	24	82.8	205	10	US-09-945-825-5	Sequence 5, Appli
39	24	82.8	299	9	US-09-769-787-82	Sequence 82, Appl
40	24	82.8	319	10	US-09-739-451-2	Sequence 2, Appli
41	24	82.8	320	10	US-09-739-451-10	Sequence 10, Appl
42	24	82.8	326	9	US-09-947-206-10	Sequence 10, Appl
43	24	82.8	326	10	US-09-016-750C-10	Sequence 10, Appl
44	24	82.8	326	10	US-09-016-869B-10	Sequence 10, Appl
45	24	82.8	331	9	US-09-992-598-209	Sequence 209, App

ALIGNMENTS

RESULT 1
US-09-910-582B-4
; Sequence 4, Application US/09910582B
; Publication No. US20030045476A1
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Mackenna, Deidre A.
; TITLE OF INVENTION: Heart Homing Conjugates
; FILE REFERENCE: P-LJ 4857
; CURRENT APPLICATION NUMBER: US/09/910,582B
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 09/326,718
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; TYPE: PRT
; LENGTH: 7
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-09-910-582B-4

Query Match 100.0%; Score 29; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVLVTSS 7
|||||||
Db 1 VVLVTSS 7

RESULT 2
US-09-782-650-3
; Sequence 3, Application US/09782650
; Patent No. US20020019350A1
; GENERAL INFORMATION:
; APPLICANT: Levine, Arnold J.
; APPLICANT: Mittlerer, Artur
; APPLICANT: Falkner, Falko-Gunter
; APPLICANT: Schefflinger, Friedrich
; APPLICANT: Dornier, Friedrich

Wed Jul 16 13:41:36 2003

APPLICANT: Edwards Lifesciences Corporation
TITLE OF INVENTION: Targeted Angiogenesis
FILE REFERENCE: 20553D-000611US
CURRENT APPLICATION NUMBER: US/09/782,650
CURRENT FILING DATE: 2001-02-19
PRIOR APPLICATION NUMBER: US 09/324,079
PRIOR FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: US 09/327,045
PRIOR FILING DATE: 1999-06-07
PRIOR APPLICATION NUMBER: PCT/US00/14988
PRIOR FILING DATE: 2000-05-31
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:targeting
OTHER INFORMATION: molecule

US-09-782-650-3
Query Match 100.0%; Score 29; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVLTSS 7
|||||||
Db 1 VLVLTSS 7

RESULT 3
US-10-128-714-3581
Sequence 3581, Application US/10128714
Publication No. US20030119013A1
GENERAL INFORMATION:
APPLICANT: Jiang, Bo
APPLICANT: Hu, Wenqi
APPLICANT: Tishkoff, Daniel
APPLICANT: Zamudio, Carlos
APPLICANT: Eroshkin, Alexey M
APPLICANT: Lemieux, Sebastien M
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and Methods of Use
FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128,714
CURRENT FILING DATE: 2002-04-23
PRIOR APPLICATION NUMBER: US 60/285,697
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: US 60/287,066
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/295,890
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/303,899
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/316,362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3581
LENGTH: 498
TYPE: PRT
ORGANISM: Aspergillus fumigatus
US-10-128-714-3581

Query Match 100.0%; Score 29; DB 9; Length 498;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVLTSS 7
|||||||
Db 207 VLVLTSS 213

RESULT 4
US-10-128-714-8581
Sequence 8581, Application US/10128714
Publication No. US20030119013A1
GENERAL INFORMATION:
APPLICANT: Jiang, Bo
APPLICANT: Hu, Wenqi
APPLICANT: Tishkoff, Daniel
APPLICANT: Zamudio, Carlos
APPLICANT: Eroshkin, Alexey M
APPLICANT: Lemieux, Sebastien M
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and Methods of Use
FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128,714
CURRENT FILING DATE: 2002-04-23
PRIOR APPLICATION NUMBER: US 60/285,697
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: US 60/287,066
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/295,890
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/303,899
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/316,362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8581
LENGTH: 522
TYPE: PRT
ORGANISM: Aspergillus fumigatus
US-10-128-714-8581

Query Match 100.0%; Score 29; DB 9; Length 522;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVLTSS 7
|||||||
Db 207 VLVLTSS 213

RESULT 5
US-09-738-626-5150
Sequence 5150, Application US/09738626
Publication No. US20020197603A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENO, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5150
LENGTH: 83

; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5150

Query Match 89.7%; Score 26; DB 9; Length 83;
Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
|:|||||
Db 43 VLVLTSA 49

RESULT 6

US-09-738-626-5345
; Sequence 5345, Application US/09738626
; Publication No. US20020197605A1

; GENERAL INFORMATION:

; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIO
; APPLICANT: OCHIALI, KEIKO
; APPLICANT: YOKOL, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patentin ver. 3.0
; SEQ ID NO 5345
; LENGTH: 188

; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum

US-09-738-626-5345

Query Match 89.7%; Score 26; DB 9; Length 188;
Best Local Similarity 57.1%; Pred. No. 1.9e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
|:|||||
Db 16 ILLITSS 22

RESULT 7

US-10-344-440-1
; Sequence 1, Application US/10344440
; Publication No. US20030131378A1

; GENERAL INFORMATION:

; APPLICANT: Arjoan, Raffi
; TITLE OF INVENTION: METHODS FOR BLOCKING RESISTANCE TO BT TOXINS IN INSECTS AND NEMATODES
; FILE REFERENCE: 6627-PA1023
; CURRENT APPLICATION NUMBER: US/10/344,440
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: 60/224,941
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: PCT/US01/41687
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 310

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-344-440-1

Query Match 89.7%; Score 26; DB 9; Length 310;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
|:|||||
Db 60 VLLVTSS 66

RESULT 8

US-10-060-230-24
; Sequence 24, Application US/10060230
; Patent No. US20020173014A1

; GENERAL INFORMATION:

; APPLICANT: HILTUNEN, Kalervo
; TITLE OF INVENTION: A method to control cellular (3R)-hydroxyacyl-CoA
; TITLE OF INVENTION: esters, precursor molecules for polyhydroxyalkanoate
; TITLE OF INVENTION: synthesis in genetically modified organisms
; FILE REFERENCE: 0365-0528P
; CURRENT APPLICATION NUMBER: US/10/060,230
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 19991667
; PRIOR FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 24
; LENGTH: 736

; TYPE: PRT

; ORGANISM: Homo sapiens
US-10-060-230-24

Query Match 89.7%; Score 26; DB 9; Length 736;
Best Local Similarity 85.7%; Pred. No. 8.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
|:|||||
Db 11 VLVLTSA 17

RESULT 9

US-10-205-823-232
; Sequence 232, Application US/10205823
; Publication No. US20030108963A1

; GENERAL INFORMATION:

; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gornatcheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746

Query Match	Best Local Similarity	Score	DB	Length	Indels	Mismatches	Gaps
Query Match	Best Local Similarity	89.7%	Score 26;	DB 9;	Length 874;	Indels 0;	Gaps 0;
Matches 5;	Conservative	71.4%;	Pred. No. 9.9e+02;	Mismatches 0;	Indels 0;	Gaps 0;	
SEQ ID NO 232	1 VLVITSS 7						
LENGTH: 874	102 VLVITSS 108						
TYPE: PRT							
ORGANISM: Homo sapiens							
US-10-205-823-232							
US-09-815-242-12739							
Sequence 12739, Application US/09815242							
Patent No. US20020061569A1							
GENERAL INFORMATION:							
APPLICANT: Haselbeck, Robert							
APPLICANT: Ohlsen, Karl L.							
APPLICANT: Zyskind, Judith W.							
APPLICANT: Wall, Daniel							
APPLICANT: Trawick, John D.							
APPLICANT: Carr, Grant J.							
APPLICANT: Yamamoto, Robert T.							
APPLICANT: Xu, H. Howard							
TITLE OF INVENTION: Identification of Essential Genes in							
TITLE OF INVENTION: Prokaryotes							
FILE REFERENCE: ELITRA.011A							
CURRENT APPLICATION NUMBER: US/09/815,242							
CURRENT FILING DATE: 2001-03-21							
PRIOR APPLICATION NUMBER: 60/191,078							
PRIOR FILING DATE: 2000-03-21							
PRIOR APPLICATION NUMBER: 60/206,848							
PRIOR FILING DATE: 2000-05-23							
PRIOR APPLICATION NUMBER: 60/207,727							
PRIOR FILING DATE: 2000-05-26							
PRIOR APPLICATION NUMBER: 60/242,578							
PRIOR FILING DATE: 2000-10-23							
PRIOR APPLICATION NUMBER: 60/253,625							
PRIOR FILING DATE: 2000-11-27							
PRIOR APPLICATION NUMBER: 60/257,931							
PRIOR FILING DATE: 2000-12-22							
PRIOR APPLICATION NUMBER: 60/269,308							
PRIOR FILING DATE: 2001-02-16							
NUMBER OF SEQ ID NOS: 14110							
SOFTWARE: FastSeq for Windows Version 4.0							
SEQ ID NO 12739							
LENGTH: 243							
TYPE: PRT							
ORGANISM: Staphylococcus aureus							
US-09-815-242-12739							
Query Match	Best Local Similarity	86.28;	Score 25;	DB 10;	Length 243;	Indels 0;	Gaps 0;
Matches 6;	Conservative	100.0%;	Pred. No. 4.2e+02;	Mismatches 0;	Indels 0;	Gaps 0;	
SEQ ID NO 12739	1 VLVITS 6						
LENGTH: 243	71 VLVITS 76						
TYPE: PRT							
ORGANISM: Staphylococcus aureus							
US-09-815-242-12739							
US-09-815-242-12739							
Sequence 12739, Application US/09815242							
Patent No. US20020061569A1							
GENERAL INFORMATION:							
APPLICANT: Haselbeck, Robert							
APPLICANT: Ohlsen, Karl L.							
APPLICANT: Zyskind, Judith W.							
APPLICANT: Wall, Daniel							
APPLICANT: Trawick, John D.							
APPLICANT: Carr, Grant J.							
APPLICANT: Yamamoto, Robert T.							
APPLICANT: Xu, H. Howard							
TITLE OF INVENTION: Identification of Essential Genes in				</			

; ORGANISM: Streptomyces avermitilis
US-10-156-761-10426

Query Match 86.2%; Score 25; DB 9; Length 272;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLVTS 7
 |||||
Db 17 VLVTS 22

RESULT 13

US-10-127-032-115
; Sequence 115, Application US/10127032
; Publication No. US20030113742A1
; GENERAL INFORMATION:
; APPLICANT: Whiteley, Marvin
; APPLICANT: Bangera, M. Gita
; APPLICANT: Lory, Stephen
; APPLICANT: Greenberg, Everett Peter
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF
; TITLE OF INVENTION: BIOFILM FORMATION
; FILE REFERENCE: UIZ-070CP
; CURRENT APPLICATION NUMBER: US/10/127,032
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/285,190
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/344,142
; PRIOR FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-127-032-115

Query Match 86.2%; Score 25; DB 9; Length 295;
Best Local Similarity 71.4%; Pred. No. 5.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVTS 7
 |||||
Db 167 ILVTS 173

RESULT 14

US-09-861-451A-36
; Sequence 36, Application US/09861451A
; Patent No. US20020068289A1
; GENERAL INFORMATION:
; APPLICANT: Commonwealth Scientific & Industrial Research Orga
; TITLE OF INVENTION: Methods of Identifying Antigen Gene Sequences
; FILE REFERENCE: FF34033/01
; CURRENT APPLICATION NUMBER: US/09/861,451A
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: Pp7273
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Deduced protein
; OTHER INFORMATION: sequence from clone PAD784
US-09-861-451A-36

Query Match 86.2%; Score 25; DB 10; Length 299;
Best Local Similarity 85.7%; Pred. No. 5.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLVTS 7
 |||||
Db 281 VLVTS 287

RESULT 15

US-10-156-761-9520
; Sequence 9520, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9520
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9520

Query Match 86.2%; Score 25; DB 9; Length 343;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VLVTS 7
 |||||
Db 83 VLVTS 88

Search completed: July 16, 2003, 13:35:58
Job time : 15.3171 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:04:48 ; Search time 4.78049 Seconds
(without alignments)
60.733 Million cell updates/sec

Title: US-09-910-582B-4
Perfect score: 29
Sequence: 1 VVLVTSS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	27	93.1	1061	1 CYPD_BACSU	008394 bacillus su
2	26	89.7	258	1 Y033_BPHPI	P51737 bacterioph
3	26	89.7	297	1 B3G5_PANTR	Q9n295 p beta-1,3-
4	26	89.7	300	1 B3G5_GORGO	Q9n293 g beta-1,3-
5	26	89.7	301	1 B3G5_PANPA	Q9n294 p beta-1,3-
6	26	89.7	308	1 B3G5_MOUSE	Q9j167 m beta-1,3-
7	26	89.7	310	1 B3G5_HUMAN	Q9y2c3 h beta-1,3-
8	26	89.7	453	1 YAW6_SCHPO	Q10181 schizosacch
9	26	89.7	490	1 DNAA_CAUCR	P35887 caulobacter
10	26	89.7	631	1 XRC1_MOUSE	Q60596 mus musculu
11	26	89.7	633	1 XRC1_HUMAN	P18887 homo sapien
12	26	89.7	648	1 WHN_HUMAN	Q15353 homo sapien
13	26	89.7	648	1 WHN_MOUSE	Q61575 mus musculu
14	26	89.7	870	1 COPG_CAEEL	Q22498 caenorhabdi
15	26	89.7	871	1 CPD2_HUMAN	Q9ubf2 homo sapien
16	26	89.7	871	1 CPD2_MOUSE	Q9ubk3 mus musculu
17	26	89.7	874	1 COPG_BOVIN	P53620 bos taurus
18	26	89.7	874	1 COPG_HUMAN	Q9y678 homo sapien
19	25	86.2	129	1 EXBD_HAEDU	Q51809 haemophilus
20	25	86.2	145	1 EXBD_PASHA	P72203 pasteurella
21	25	86.2	147	1 EXBD_HAEIN	P43009 haemophilus
22	25	86.2	201	1 Y488_MYCTU	Q11154 mycobacteri
23	25	86.2	210	1 ACPD_RHIL0	Q981f8 rhizobium i
24	25	86.2	223	1 GTX1_TOBAC	Q03662 nicotiana t
25	25	86.2	223	1 GTX3_TOBAC	Q03664 nicotiana t
26	25	86.2	229	1 COX2_PISOC	P25002 pisaster oc
27	25	86.2	251	1 PHOB_CAMJE	Q9pj29 campylobact
28	25	86.2	399	1 PHOC_PSEAE	P43336 pseudomonas
29	25	86.2	457	1 DNAA_HELPJ	Q9zj96 helicobacte
30	25	86.2	457	1 DNAA_HELPY	Q26057 helicobacte
31	25	86.2	498	1 TLCA_RICPR	P19568 rickettsia
32	25	86.2	524	1 CFTR_RAT	P34158 rattus norv
33	25	86.2	526	1 C4E2_DROME	Q27606 drosophila

RESULT 1
ID CYPD_BACSU STANDARD; PRT; 1061 AA.
AC 008394;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable bifunctional P-450:NADPH-P450 reductase 1 [Includes:
DE Cytochrome P450 102 (EC 1.14.14.1); NADPH-cytochrome P450 reductase
DE (EC 1.6.2.4)].
GN CYPD OR CYP102A2.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=97453479; PubMed=9308178;
RA Sorokin A., Bolotin A., Purnelle B., Hilbert H., Lauber J.,
RA Duesterhoeft A., Ehrlich S.D.;
RT "Sequence of the Bacillus subtilis genome region in the vicinity of
the lev operon reveals two new extracytoplasmic function RNA
polymerase sigma factors SigV and SigZ.";
RT Microbiology 143:2939-2943(1997).
RL [2]
RN SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
RA Ghim S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Haeck J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaert-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapiou A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale F.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone E.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Taconis E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,

ALIGNMENTS

34 25 86.2 565 1 FUTI_PEA
35 25 86.2 571 1 ILVI_BUCAI
36 25 86.2 571 1 ILVI_BUCAP
37 25 86.2 571 1 ILVI_BACSU
38 25 86.2 574 1 ILVI_ECOLI
39 25 86.2 574 1 ILVI_SALTY
40 25 86.2 669 1 ILVB_SCHPO
41 25 86.2 687 1 ILVB_YEAST
42 25 86.2 905 1 COPG_SCHPO
43 25 86.2 932 1 K6PL_KLULA
44 25 86.2 1035 1 CC68_YEAST
45 25 86.2 2747 1 FAF_DROME

Q9m5q1 pismu savi
P57321 buchnera ap
O85293 buchnera ap
Q04789 bacillus su
P00893 escherichia
P40811 salmonella
P36620 schizosacch
P07342 saccharomyc
P87140 schizosacch
Q03215 kluyveromyc
P32558 saccharomyc
P55824 drosophila

RT The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.";

RL Nature 390:249-256(1997).

CC -!- FUNCTION: FUNCTIONS AS A FATTY ACID MONOOXYGENASE. THE REDUCTASE DOMAIN IS REQUIRED FOR ELECTRON TRANSFER FROM NADP TO CYTOCHROME P450 (BY SIMILARITY).

CC -!- CATALYTIC ACTIVITY: NADPH + 2 ferricytochrome = NADP(+) + 2 ferrocycytochrome.

CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH + oxidized flavoprotein + H(2)O.

CC -!- COFACTOR: FAD AND FMN (BY SIMILARITY).

CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTOCHROME P450 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: D87979; BAA20123.1; .

CC EMBL: 299107; CAB12544.1; .

CC HSSP: P14779; LJPZ.

CC Subtilist; BG12871; cypD.

CC InterPro: IPR001128; Cytochrome_P450.

CC InterPro: IPR003097; FAD_binding.

CC InterPro: IPR001709; FPN_cyt_reductse.

CC InterPro: IPR001226; Flavodoxin.

CC InterPro: IPR001433; Oxred_FAD/NAD(P).

CC Pfam: PF00057; P450; 1.

CC Pfam: PF00175; NAD_binding; 1.

CC Pfam: PF00258; flavodoxin; 1.

CC Pfam: PF00667; FAD_binding; 1.

CC PRINTS: PR00371; FPNCR.

CC PRINTS: PR00385; P450.

CC PROSITE: PS00086; CYTOCHROME_P450; 1.

CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;

CC Multifunctional enzyme; FMN; FAD; Flavoprotein; NADP;

CC Complete proteome.

CC DOMAIN 1 474 CYTOCHROME P450.

CC DOMAIN 475 1061 NADPH-P-450 REDUCTASE.

CC BINDING 403 403 HEME (BY SIMILARITY).

CC SEQUENCE 1061 AA; 119467 MW; 7915DACC20578978 CRC64;

Query Match 93.1%; Score 27; DB 1; Length 1061;

Best Local Similarity 85.7%; Pred. No. 1.4e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVLVTSS 7

Db 540 VVIWSS 546

RESULT 2

Y033 BPHP1 STANDARD; PRT; 258 AA.

AC P51737;

DT 01-OCT-1996 (Rel. 34, Created).

DT 01-OCT-1996 (Rel. 34, Last sequence update).

DT 16-OCT-2001 (Rel. 40, Last annotation update).

DE Hypothetical 28.3 kDa protein in LYS 3' region (ORF33).

OS Bacteriophage HP1.

OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;

OC unclassified Myoviridae.

OC NCBI_TaxID=10690;

OX [1]

RN SEQUENCE FROM N.A.

RP STRAIN=HP1C1;

RC MEDLINE=96279738; PubMed=8710508;

RA Esposito D., Fitzmaurice W.P., Benjamin R.C., Goodman S.D., Waldman A.S., Scoeca J.J.;

RT The complete nucleotide sequence of bacteriophage HP1 DNA.";

RL Nucleic Acids Res. 24:2360-2368(1996).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: U24159; AAB09220.1; .

CC Hypothetical protein.

KW Hypothetical protein.

SQ SEQUENCE 258 AA; 28413 MW; 45A9235B70672E42 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 258;

Best Local Similarity 71.4%; Pred. No. 54;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVLVTSS 7

Db 87 VIMVTSS 93

RESULT 3

B3G5_PANTR STANDARD; PRT; 297 AA.

ID B3G5_PANTR

AC Q9N295;

DT 16-OCT-2001 (Rel. 40, Created).

DT 16-OCT-2001 (Rel. 40, Last sequence update).

DT 15-JUN-2002 (Rel. 41, Last annotation update).

DE Beta-1,3-galactosyltransferase 5 (EC 2.4.1.-) (Beta-1,3-GalTase 5)

DE (Beta3gal-T5) (b3Gal-T5) (UDP-galactose-beta-N-acetylglucosamine beta-1,3-galactosyltransferase 5) (UDP-gal-beta-GlcNAc beta-1,3-galactosyltransferase 5) (Beta-3-Gx-T5) (Fragment).

GN B3GALT5.

OS Pan troglodytes (Chimpanzee).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.

OX NCBI_TaxID=9598;

RN [1]

RP SEQUENCE FROM N.A.

RA Liu Y., Saitou N.;

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: CATALYZES THE TRANSFER OF GAL TO GLCNAC-BASED ACCEPTORS WITH A PREFERENCE FOR THE CORE3 O-LINKED GLYCAN GLCNAC(BETA1,3)GALNAC STRUCTURE. CAN USE GLYCOPOLIPID LC3CER AS AN EFFICIENT ACCEPTOR (BY SIMILARITY).

CC -!- PATHWAY: Glycosylation.

CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (potential).

CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 31.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: AB041414; BAA94499.1; .

DR InterPro: IPR002659; Galactosyl_T.

DR Pfam: PF01762; Galactosyl_T; 1.

KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;

KW Signal-anchor; Golgi stack; Multigene family.

FT DOMAIN 1 7 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 8 28 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).

FT LUMENAL, CATALYTIC (POTENTIAL).

FT DOMAIN 29 >297

FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT NON_TER 297 297

SQ SEQUENCE 297 AA; 34850 MW; BB7963250A837A28 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 297;
 Best Local Similarity 85.7%; Pred. No. 63;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVVTSS 7
 I:|||||
 Db 60 VLLVTSS 66

RESULT 4
 B3G5_GORGO
 ID B3G5_GORGO STANDARD: PRT; 300 AA.
 AC Q9N293;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Beta-1,3-galactosyltransferase 5 (EC 2.4.1.-) (Beta-1,3-GalTase 5)
 DE (Beta3Gal-T5) (b3Gal-T5) (UDP-galactose:beta-N-acetylglucosamine beta-
 1,3-galactosyltransferase 5) (UDP-Gal:beta-GlcNAc beta-1,3-
 DE galactosyltransferase 5) (Beta-3-Gx-T5) (Fragment).
 CN B3GALT5.
 OS Gorilla gorilla gorilla (Lowland gorilla).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 OX NCBI_TaxID=9595;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Liu Y., Saitou N.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE TRANSFER OF GAL TO GLCNAC-BASED ACCEPTORS
 CC WITH A PREFERENCE FOR THE CORE3 O-LINKED GLYCAN
 CC GLCNAC(BETA1,3)GALNAC STRUCTURE. CAN USE GLYCOPOLIPID LC3CER AS AN
 CC EFFICIENT ACCEPTOR (BY SIMILARITY).
 CC -!- PATHWAY: Glycosylation.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (Potential).
 CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 31.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB041416; BAA94501.1; -
 DR InterPro; IPR002659; Galactosyl_T.
 DR Pfam; PF01762; Galactosyl_T; 1.
 KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
 KW Signal-anchor; Golgi stack; Multigene family;
 FT DOMAIN 1 7
 FT TRANSMEM 8 28
 FT SIGNAL-ANCHOR (POTENTIAL).
 FT CYTOPLASMIC (POTENTIAL).
 FT LUMENAL, CATALYTIC (POTENTIAL).
 FT CARBOHYD 130 130
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 174 174
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 231 231
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT NON_TER 300 300
 SQ SEQUENCE 300 AA; 35104 MW; 1AA11692ED9F06FA CRC64;

Query Match 89.7%; Score 26; DB 1; Length 300;
 Best Local Similarity 85.7%; Pred. No. 64;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVVTSS 7
 I:|||||
 Db 60 VLLVTSS 66

RESULT 5
 B3G5_PANPA

ID B3G5_PANPA STANDARD: PRT; 301 AA.
 AC Q9N294;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Beta-1,3-galactosyltransferase 5 (EC 2.4.1.-) (Beta-1,3-GalTase 5)
 DE (Beta3Gal-T5) (b3Gal-T5) (UDP-galactose:beta-N-acetylglucosamine beta-
 1,3-galactosyltransferase 5) (UDP-Gal:beta-GlcNAc beta-1,3-
 DE galactosyltransferase 5) (Beta-3-Gx-T5) (Fragment).
 CN B3GALT5.
 OS Pan paniscus (Pygmy chimpanzee) (Bonobo).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9597;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Liu Y., Saitou N.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE TRANSFER OF GAL TO GLCNAC-BASED ACCEPTORS
 CC WITH A PREFERENCE FOR THE CORE3 O-LINKED GLYCAN
 CC GLCNAC(BETA1,3)GALNAC STRUCTURE. CAN USE GLYCOPOLIPID LC3CER AS AN
 CC EFFICIENT ACCEPTOR (BY SIMILARITY).
 CC -!- PATHWAY: Glycosylation.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (Potential).
 CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 31.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB041415; BAA94500.1; -
 DR InterPro; IPR002659; Galactosyl_T.
 DR Pfam; PF01762; Galactosyl_T; 1.
 KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
 KW Signal-anchor; Golgi stack; Multigene family;
 FT DOMAIN 1 7
 FT TRANSMEM 8 28
 FT SIGNAL-ANCHOR (POTENTIAL).
 FT CYTOPLASMIC (POTENTIAL).
 FT LUMENAL, CATALYTIC (POTENTIAL).
 FT CARBOHYD 130 130
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 174 174
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 231 231
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT NON_TER 301 301
 SQ SEQUENCE 301 AA; 35222 MW; E04CC70F41CB29FD CRC64;

Query Match 89.7%; Score 26; DB 1; Length 301;
 Best Local Similarity 85.7%; Pred. No. 64;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVVTSS 7
 I:|||||
 Db 60 VLLVTSS 66

RESULT 6
 B3G5_MOUSE
 ID B3G5_MOUSE STANDARD: PRT; 308 AA.
 AC Q9J167;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Beta-1,3-galactosyltransferase 5 (EC 2.4.1.-) (Beta-1,3-GalTase 5)
 DE (Beta3Gal-T5) (b3Gal-T5) (UDP-galactose:beta-N-acetylglucosamine beta-
 1,3-galactosyltransferase 5) (UDP-Gal:beta-GlcNAc beta-1,3-
 DE galactosyltransferase 5) (Beta-3-Gx-T5) (Stage-specific embryonic
 DE antigen-3 synthase) (SSEA-3 synthase).
 GN B3GALT5 OR B3GT5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=20390006; PubMed=10837462;
RA Zhou D., Henion T.R., Jungalwala F.B., Berger E.G., Hennet T.;
RT "The beta1,3-galactosyltransferase beta3galr-v is a stage-specific
embryonic antigen-3 (SSEA-3) synthase.";
RL J. Biol. Chem. 275:22631-22634(2000).
CC -|- FUNCTION: CATALYZES THE TRANSFER OF GAL TO GLCNAC-BASED ACCEPTORS
WITH A PREFERENCE FOR THE CORE3 O-LINKED GLYCAN
GLCNAC(BETA1,3)GALNAC STRUCTURE. CAN USE GLYCOPOLIPID LC3CER AS AN
EFFICIENT ACCEPTOR. ALSO CATALYZES THE TRANSFER OF GAL TO THE
TERMINAL GALNAC UNIT OF THE GLOBOSIDE GB4, THEREBY SYNTHESIZING
THE GLYCOPOLIPID GB5, ALSO KNOWN AS THE STAGE-SPECIFIC EMBRYONIC
ANTIGEN-3 (SSEA-3).
CC -|- PATHWAY: Glycosylation.
CC -|- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (Potential).
CC -|- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND KIDNEY.
CC -|- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 31.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF254738; AAF86241.1;
DR MGD; MGI:2136878; B3gal5.
DR InterPro: IPR002659; Galactosyl_T.
DR Pfam: PF01762; Galactosyl_T; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack; Multigene family.
FT DOMAIN 1 7
FT TRANSMEM 8 25
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT CYTOPLASMIC (POTENTIAL).
FT (POTENTIAL).
FT LUMENAL, CATALYTIC (POTENTIAL).
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 128 308
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 172 172
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 229 229
FT SEQUENCE 308 AA; 35964 MW; 789073A5178825B1 CRC64;
SQ

Query Match 89.7%; Score 26; DB 1; Length 308;
Best Local Similarity 85.7%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
D 58 VLLVTS 64

RESULT 7
B3G5_HUMAN STANDARD; PRT; 310 AA.
ID B3G5_HUMAN STANDARD; PRT; 310 AA.
AC Q9Y2C3; Q9NY96; Q9P1X7; Q9P1X6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-1,3-galactosyltransferase 5 (EC 2.4.1.-) (Beta-1,3-GalTase 5)
DE (Beta3Gal-T5) (UDP-galactose:beta-N-acetylglucosamine beta-
1,3-galactosyltransferase 5) (UDP-Gal:beta-GlcNAc beta-1,3-
galactosyltransferase 5) (Beta-3-Gx-T5).
GN B3GALT5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colorectal adenocarcinoma;

RX MEDLINE=99230269; PubMed=10212226;
RA Isshiki S., Toyayachi A., Kudo T., Nishihara S., Watanabe M.,
RA Kubota T., Kitajima M., Shiraishi N., Sasaki K., Andoh T.,
RA Narimatsu H.;
RT "Cloning, expression, and characterization of a novel UDP-
galactose:beta-N-acetylglucosamine beta1,3-galactosyltransferase
(beta3gal-T5) responsible for synthesis of type 1 chain in colorectal
and pancreatic epithelia and tumor cells derived therefrom.";
RL J. Biol. Chem. 274:12499-12507(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99337698; PubMed=10406968;
RA Zhou D., Berger E.G., Hennet T.;
RT "Molecular cloning of a human UDP-galactose:GlcNAc beta1,3GalNAC
beta1,3 galactosyltransferase gene encoding an O-linked
core3-elongation enzyme.";
RL Eur. J. Biochem. 263:571-576(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Amado M., Carneiro F., Clausen H.;
RT "Cloning and expression of two beta-1,3-galactosyltransferases:
beta3gal-T5 and beta3gal-T6.";
RN Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Ohki H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Park M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordliek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RN Nature 405:311-319(2000).
RN [5]
RP SEQUENCE OF 1-298 FROM N.A.
RA Liu Y., Saltou N.;
RN Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP REVIEW.
RX MEDLINE=20047730; PubMed=10580128;
RA Amado M., Almeida R., Schwientek T., Clausen H.;
RT "Identification and characterization of large galactosyltransferase
gene families: galactosyltransferases for all functions.";
RL Biochim. Biophys. Acta 1473:35-53(1999).
CC -|- FUNCTION: CATALYZES THE TRANSFER OF GAL TO GLCNAC-BASED ACCEPTORS
WITH A PREFERENCE FOR THE CORE3 O-LINKED GLYCAN
GLCNAC(BETA1,3)GALNAC STRUCTURE. CAN USE GLYCOPOLIPID LC3CER AS AN
EFFICIENT ACCEPTOR.
CC -|- PATHWAY: Glycosylation.
CC -|- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (Potential).
CC -|- TISSUE SPECIFICITY: EXPRESSED IN STOMACH, JEJUNUM, COLON,
PANCREAS, SMALL INTESTINE, TESTIS AND GASTROINTESTINAL AND
PANCREATIC CANCER CELL LINES. HARDLY DETECTED IN LUNG, LIVER,
ADRENAL GLAND AND PERIPHERAL BLOOD LEUKOCYTES.
CC -|- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 31.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB020337; BAA77664.1;
DR EMBL; AF145784; AAF07880.1;

```

DR EMBL: AJ006078; CAB91547.1; -
DR EMBL: AL163280; CAB90446.1; -
DR EMBL: AB041412; BAA94497.1; -
DR EMBL: AB041413; BAA94498.1; -
DR Genew: HGNC:920; B3GALT5.
DR MIM: 604066; -
DR InterPro: IPR002659; Galactosyl_T.
DR Pfam: PF01762; Galactosyl_T; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack; Multigene family.
FT DOMAIN 1 7 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 8 28 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 29 310 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 130 130 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 174 174 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 231 231 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 23 23 F -> Y (IN REF. 3).
FT CONFLICT 26 26 Y -> N (IN REF. 3).
FT CONFLICT 42 42 G -> V (IN REF. 5; BAA94497).
FT CONFLICT 85 85 M -> T (IN REF. 3 AND 5; BAA94497).
FT CONFLICT 114 114 Q -> R (IN REF. 5; BAA94497).
FT CONFLICT 234 234 K -> E (IN REF. 5; BAA94497).
FT CONFLICT 277 277 L -> R (IN REF. 5; BAA94497).
FT CONFLICT 293 297 LDYWQ -> WTTGR (IN REF. 5; BAA94498).
SQ SEQUENCE 310 AA; 36189 MW; 4DD7A19E3E648AA9 CRC64;

Query Match 89.78; Score 26; DB 1; Length 310;
Best Local Similarity 85.78; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
|:|||||
DB 60 VLLVTSS 66

RESULT 8
YAW6_SCHPO STANDARD; PRT; 453 AA.
AC Q10181;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C3F10.06c in chromosome I.
GN SPAC3F10.06C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
RA Squeros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feitwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidaigo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu A., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

```

```

RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RL Nature 415:871-880(2002).
CC -!- SIMILARITY: TO YEAST RIT1.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z69369; CAA93304.1; -
KW Hypothetical protein.
SQ SEQUENCE 453 AA; 50850 MW; ADBFDA50D7ACDEC8 CRC64;

Query Match 89.78; Score 26; DB 1; Length 453;
Best Local Similarity 85.78; Pred. No. 99;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
|:|||||
DB 210 VLVLTSS 216

RESULT 9
DNAA_CAUCR STANDARD; PRT; 490 AA.
ID DNAA_CAUCR
AC P35887;
DT 01-JUN-1994 (Rel. 29, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chromosomal replication initiator protein dnaA.
GN DNAA OR CC0008.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Cbl5N / NA1000;
RX MEDLINE=94117375; PubMed=8288535;
RA Zweiger G., Shapiro L.;
RT "Expression of Caulobacter dnaA as a function of the cell cycle.";
RL J. Bacteriol. 176:401-408(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / Cbl5;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocha I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE INITIATION AND REGULATION
CC OF CHROMOSOMAL REPLICATION. BINDS TO THE ORIGIN OF REPLICATION;
CC IT BINDS SPECIFICALLY DOUBLE-STRANDED DNA AT A 9 BP CONSENSUS
CC (DNAA BOX): 5'-TTATC(C/A)A(C/A)-3'. DNAA BINDS TO ATP AND TO
CC ACIDIC PHOSPHOLIPIDS.
CC -!- SIMILARITY: BELONGS TO THE DNAA FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: U01667; AAA18927.1; -
DR EMBL; AE005675; AAK21996.1; -
DR PIR; B36947; B36947.
DR TIGR; CC0008; -
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001357; Bac_DnaA.
DR Pfam; PF00308; bac_dnaa; 1.
DR PRINTS; PR00051; DAAA.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00362; DnaA; 1.
DR PROSITE; PS01008; DAAA; 1.
KW DNA replication; DNA-binding; ATP-binding; Complete proteome.
FT NP_BIND 189 196 H->L (IN REF. 1).
FT CONFLICT 334 334
FT SEQUENCE 490 AA; 53883 MW; 33BD9484340B5DC1 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 490;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
Db 60 VLVLTST 66

RESULT 10
XRCL_MOUSE STANDARD; PRT; 631 AA.
AC Q60596;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA-repair protein XRCC1.
GN XRCC1 OR XRC-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95048367; PubMed=7959765;
RA Brookman K.W., Tebbes R.S., Allen S.A., Tucker J.D., Swiger R.R.,
RA Lamerdin J.E., Carrano A.V., Thompson L.H.;
RA "Isolation and characterization of mouse Xrcc-1, a DNA repair gene
affecting ligation."
RL Genomics 22:180-188(1994).
CC -!- FUNCTION: CORRECTS DEFECTIVE DNA STRAND-BREAK REPAIR AND SISTER
CC CHROMATID EXCHANGE FOLLOWING TREATMENT WITH IONIZING RADIATION
CC AND ALKYLATING AGENTS.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: SOME TO S. POMBE RAD4/CUT5.
CC -!- SIMILARITY: CONTAINS 2 BCT DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U02887; AAA93115.1; -
DR MGD; MGI:99137; Xrccl.
DR InterPro; IPR001357; BRC1.
DR InterPro; IPR002706; XRCC1_N.
DR Pfam; PF00533; BRC1; 2.
DR Pfam; PF01834; XRCC1_N; 1.

ProDom: PD023136; Xrccl_N; 1.
DR SMART; SM00292; BRC1; 2.
DR PROSITE; PS0172; BRC1; 2.
KW DNA repair; Nuclear protein; Repeat.
FT DOMAIN 315 403
FT DOMAIN 536 627
FT SEQUENCE 631 AA; 69003 MW; 7F1868BFEB2A3C68 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 631;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
Db 88 VLVLTSS 94

RESULT 11
XRCL_HUMAN STANDARD; PRT; 633 AA.
ID XRCL_HUMAN
AC P18887; O9HCB1;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA-repair protein XRCC1.
GN XRCC1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91061722; PubMed=2247054;
RA Thompson L.H., Brookman K.W., Jones N.J., Allen S.A., Carrano A.V.;
RA "Molecular cloning of the human XRCC1 gene, which corrects defective
DNA strand break repair and sister chromatid exchange."
RL Mol. Cell. Biol. 10:6160-6171(1990).
RN [2]
RP SEQUENCE OF 170-633 FROM N.A.
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Ganes J.,
RA Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,
RA Andreise T., Trankheim M., Attix C., Amico-Keller G., Coefield J.,
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RA Olsen A.S., Carrano A.V.;
RA "Sequence analysis of 19q13.2."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

INTERACTIONS
RX MEDLINE=21097725; PubMed=11163244;
RA Whitehouse C.J., Taylor R.M., Thistlethwaite A., Zhang H.,
RA Karini-Busheri F., Lasko D.D., Weinfield M., Caldecott K.W.;
RA "XRCC1 stimulates human polynucleotide kinase activity at damaged DNA
termini and accelerates DNA single-strand break repair."
RL Cell 104:107-117(2001).
RN [4]
RP VARIANTS TRP-194; HIS-280 AND GLN-399.
RX MEDLINE=98143538; PubMed=9485007;
RA Shen M.R., Jones I.M., Mohrenweiser H.;
RA "Nonconservative amino acid substitution variants exist at polymorphic
frequency in DNA repair genes in healthy humans."
RL Cancer Res. 58:604-608(1998).
RN [5]
RP VARIANT GLN-399.
RX MEDLINE=20247096; PubMed=10783319;
RA Duell E.J., Wiencke J.K., Cheng T.J., Varkonyi A., Zuo Z.F.,
RA Ashok T.D., Mark E.J., Wain J.C., Christiani D.C., Kelsey K.T.;
RA "Polymorphisms in the DNA repair genes XRCC1 and ERCC2 and biomarkers
of DNA damage in human blood mononuclear cells."
RL Carcinogenesis 21:965-971(2000).
RN [6]

```
RP VARIANT GLN-399.
RX MEDLINE-21642060; PubMed-11782372;
RA Nelson H.H., Kelsey K.T., Mott L.A., Karagas M.R.;
RT "The XRCC1 Arg399Gln polymorphism, sunburn, and non-melanoma skin
RT cancer: evidence of gene-environment interaction.";
RL Cancer Res. 62:152-155(2002).
CC -1- FUNCTION: CORRECTS DEFECTIVE DNA STRAND-BREAK REPAIR AND SISTER
CC CHROMATID EXCHANGE FOLLOWING TREATMENT WITH IONIZING RADIATION
CC AND ALKYLATING AGENTS.
CC -1- SUBUNIT: Interacts with polynucleotide kinase (PNK), DNA
CC polymerase-beta (POLB) and DNA ligase III (LIG3).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- POLYMORPHISM: Carriers of the polymorphic Gln-399 allele may be at
CC greater risk for tobacco- and age-related DNA damage.
CC -1- SIMILARITY: SOME, TO S.POMBE RAD4/CUT5.
CC -1- SIMILARITY: CONTAINS 2 BRC2 DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M36089; AAAG3270.1; -.
DR EMBL; AC018758; AAG09061.1; -.
DR PIR; A36353; A36353.
DR Genew; HGNC:12828; XRCC1.
DR MIM; 194360; -.
DR InterPro; IPR001357; BRC2.
DR InterPro; IPR002706; Xrcc1_N.
DR Pfam; PF00533; BRC2; 2.
DR Pfam; PF01834; XRCC1_N; 1.
DR ProDom; PD0231136; Xrcc1_N; 1.
DR SMART; SM00292; BRC2; 2.
DR PROSITE; PS01172; BRC2; 2.
DR PROSITE; PS01172; BRC2; 2.
DR DNA repair; Nuclear protein; Repeat; Polymorphism.
FT DOMAIN 315 403
FT BRC2 1.
FT DOMAIN 538 629
FT BRC2 2.
FT VARIANT 194 194
FT R -> W.
FT /FTID-VAR_013400.
FT VARIANT 280 280
FT R -> H.
FT /FTID-VAR_013401.
FT VARIANT 399 399
FT R -> Q.
FT /FTID-VAR_011487.
FT CONFLICT 576 576
FT Y -> N (IN REF. 2).
SQ SEQUENCE 633 AA; 69525 MW; 30CC2421345ABFC2 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 633;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
Db 86 VLVLTSS 92

RESULT 12
WHN_HUMAN
ID WHN_HUMAN STANDARD; PRT; 648 AA.
AC O15353; O15352;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription factor winged-helix nude (Forkhead box protein N1).
GN WHN OR FOXN1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT VAL-283.
```

```
RC TISSUE=Thymus;
RX MEDLINE-98025083; PubMed-9321431;
RA Schorpp M., Hofmann M., Dear T.N., Boehm T.;
RT "Characterization of mouse and human nude genes.";
RL Immunogenetics 46:509-515(1997).
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR INVOLVED IN DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THYMUS.
CC -1- SIMILARITY: CONTAINS 1 FORK-HEAD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y11741; CAA72417.1; -.
DR EMBL; Y11742; CAA72417.1; JOINED.
DR EMBL; Y11743; CAA72417.1; JOINED.
DR EMBL; Y11744; CAA72417.1; JOINED.
DR EMBL; Y11745; CAA72417.1; JOINED.
DR EMBL; Y11746; CAA72417.1; JOINED.
DR EMBL; Y11739; CAA72416.1; -.
DR HSSP; Q63245; 2HFH.
DR TRANSFAC; T02525; -.
DR Genew; HGNC:12765; WHN.
DR MIM; 600838; -.
DR InterPro; IPR001766; TF_Fork_head.
DR Pfam; PF00250; Fork_head; 1.
DR PRINTS; PR00053; FORKHEAD.
DR ProDom; PD000425; TF_Fork_head; 1.
DR SMART; SM00339; FH; 1.
DR PROSITE; PS00657; FORK_HEAD_1; FALSE_NEG.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS0039; FORK_HEAD_3; 1.
DR DNA-binding; Nuclear protein; Transcription regulation;
KW Developmental protein; Polymorphism.
FT DNA_BIND 271 367
FT FORK-HEAD.
FT VARIANT 283 283
FT A -> V.
FT /FTID-VAR_010376.
SQ SEQUENCE 648 AA; 68924 MW; F32F4C95627E60DE CRC64;

Query Match 89.7%; Score 26; DB 1; Length 648;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
Db 559 VLVLTSS 565

RESULT 13
WHN_MOUSE
ID WHN_MOUSE STANDARD; PRT; 648 AA.
AC Q61575;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription factor winged-helix nude (Forkhead box protein N1)
DE (Hepatocyte nuclear factor 3 forkhead homolog 1) (HNF-3/forkhead
DE homolog 1) (HFN-1).
GN WHN OR FOXN1 OR HFN1 OR FKHL9.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Skin;
RX MEDLINE-95059376; PubMed-7969402;
RA Nehls M., Pfeiffer D., Schorpp M., Hedrich H., Boehm T.;
```



```
RT "New member of the winged-helix protein family disrupted in mouse and
RT rat nude mutations."
RL Nature 372:103-107(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=BA16/C; TISSUE=Skin, and Thymus;
RX MEDLINE=98025083; PubMed=9321431;
RT Schorpp M., Hofmann M., Dear T.N., Boehm T.;
RT "Characterization of mouse and human nude genes.";
RL Immunogenetics 46:509-515(1997).
CC -1- FUNCTION: TRANSCRIPTIONAL REGULATOR INVOLVED IN DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THYMUS AND SKIN.
CC -1- SIMILARITY: CONTAINS 1 FORK-HEAD DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X81593; CAA57279.1; .
DR TRANSFAC; T01835; .
DR HSP; O63245; 2FH.
DR MGD; MGI:102949; Foxl1.
DR InterPro; IPR001766; TF_Fork_head.
DR Pfam; PF00250; Fork_head; 1.
DR PRINTS; PR00053; FORKHEAD.
DR ProDom; PD000425; TF_Fork_head; 1.
DR SMART; SM00339; FH; 1.
DR PROSITE; PS00657; FORK_HEAD_1; FALSE_NEG.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS00639; FORK_HEAD_3; 1.
DR DNA-binding; Nuclear protein; Transcription regulation;
KW Developmental protein.
FT DNA_BIND 271 367 FORK-HEAD.
FT DOMAIN 573 576 POLY-PRO.
SQ SEQUENCE 648 AA; 69244 MW; 7A75411BDE2B23E5 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 648;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
DB 559 LVLVTSS 565

RESULT 14
COPG_CAEEL STANDARD; PRT; 870 AA.
AC Q22498;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable coatmer gamma subunit (Gamma-coat protein) (Gamma-COP).
GN T14G10.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wild A.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THE COATOMER IS A CYTOSOLIC PROTEIN COMPLEX THAT BINDS
CC TO DILYSINE MOTIFS AND REVERSIBLY ASSOCIATES WITH GOLGI NON-
CC CLATHRIN-COATED VESICLES, WHICH FURTHER MEDIATE BIOSYNTHETIC
CC PROTEIN TRANSPORT FROM THE ER, VIA THE GOLGI UP TO THE TRANS GOLGI
CC NETWORK. COATOMER COMPLEX IS REQUIRED FOR BUDDING FROM GOLGI
CC -----
CC MEMBRANES, AND IS ESSENTIAL FOR THE RETROGRADE GOLGI-TO-ER
CC TRANSPORT OF DILYSINE-TAGGED PROTEINS. IN MAMMALS, THE COATOMER
CC CAN ONLY BE RECRUITED BY MEMBRANES ASSOCIATED TO ADP-RIBOSYLATION
CC FACTORS (ARFs), WHICH ARE SMALL GTP-BINDING PROTEINS; THE COMPLEX
CC ALSO INFLUENCES THE GOLGI STRUCTURAL INTEGRITY, AS WELL AS THE
CC PROCESSING, ACTIVITY, AND ENDOCYTIC RECYCLING OF LDL RECEPTORS (BY
CC SIMILARITY).
CC -1- SUBUNIT: OLIGOMERIC COMPLEX. BINDS TO CDC42.
```

```
CC MEMBRANES, AND IS ESSENTIAL FOR THE RETROGRADE GOLGI-TO-ER
CC TRANSPORT OF DILYSINE-TAGGED PROTEINS (BY SIMILARITY).
CC -1- SUBUNIT: OLIGOMERIC COMPLEX THAT CONSISTS OF AT LEAST THE ALPHA,
CC BETA, BETA', GAMMA, DELTA, EPSILON AND ZETA SUBUNITS (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: THE COATOMER IS CYTOPLASMIC OR POLYMERIZED
CC ON THE CYTOPLASMIC SIDE OF THE GOLGI, AS WELL AS ON THE
CC VESICLES/BUDS ORIGINATING FROM IT (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE COPG FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z68880; CAA93095.1; .
DR WormPep; T14G10.5; CE06451.
DR InterPro; IPR002553; Adaptin_N.
DR Pfam; PF01602; Adaptin_N; 1.
DR Transprot; Protein transport; Golgi stack; Membrane.
KW SEQUENCE 870 AA; 96302 MW; 24CCA86160A60049 CRC64;
SQ SEQUENCE 870 AA; 96302 MW; 24CCA86160A60049 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 870;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
DB 98 VIIVTSS 104

RESULT 15
COPG2_HUMAN STANDARD; PRT; 871 AA.
ID COPG2_HUMAN
AC Q9UBF2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coatomer gamma-2 subunit (Gamma-2 coat protein) (Gamma-2 COP).
GN COPG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Whitney J.A., Godzich M., Kreis T.E.;
RT "Newly identified coatomer subunits reveal multiple copy complexes
RT acting in the early secretory pathway.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20025747; PubMed=10556286;
RA Blagitzko N., Schulz U., Schinzel A.A., Ropers H.-H., Kalscheuer V.M.;
RT "Gamma2-COP, a novel imprinted gene on chromosome 7q32, defines a new
RT imprinting cluster in the human genome.";
RL Hum. Mol. Genet. 8:2387-2396(1999).
CC -1- FUNCTION: THE COATOMER IS A CYTOSOLIC PROTEIN COMPLEX THAT BINDS
CC TO DILYSINE MOTIFS AND REVERSIBLY ASSOCIATES WITH GOLGI NON-
CC CLATHRIN-COATED VESICLES, WHICH FURTHER MEDIATE BIOSYNTHETIC
CC PROTEIN TRANSPORT FROM THE ER, VIA THE GOLGI UP TO THE TRANS GOLGI
CC NETWORK. COATOMER COMPLEX IS REQUIRED FOR BUDDING FROM GOLGI
CC -----
CC MEMBRANES, AND IS ESSENTIAL FOR THE RETROGRADE GOLGI-TO-ER
CC TRANSPORT OF DILYSINE-TAGGED PROTEINS. IN MAMMALS, THE COATOMER
CC CAN ONLY BE RECRUITED BY MEMBRANES ASSOCIATED TO ADP-RIBOSYLATION
CC FACTORS (ARFs), WHICH ARE SMALL GTP-BINDING PROTEINS; THE COMPLEX
CC ALSO INFLUENCES THE GOLGI STRUCTURAL INTEGRITY, AS WELL AS THE
CC PROCESSING, ACTIVITY, AND ENDOCYTIC RECYCLING OF LDL RECEPTORS (BY
CC SIMILARITY).
CC -1- SUBUNIT: OLIGOMERIC COMPLEX. BINDS TO CDC42.
```


CC -1- SUBCELLULAR LOCATION: THE COMOTER IS CYTOPLASMIC OR POLYMERIZED
CC ON THE CYTOPLASMIC SIDE OF THE GOLGI, AS WELL AS ON THE
CC VESICLES/BUDS ORIGINATING FROM IT (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE COPG FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF207598; AAF19433.1; -
CC EMBL; AF157833; AAF14271.1; -
CC Genew; HGNC:2237; COPG2.
CC MIM; 604355; -
CC InterPro; IPR002553; Adaptin_N.
CC Pfam; PF01602; Adaptin_N; 1.
CC Transport; Protein transport; Golgi stack; Membrane.
CC KW SEQUENCE 871 AA; 97621 MW; D361A0E74069CBEB CRC64;
CC
CC Query Match 89.7%; Score 26; DB 1; Length 871;
CC Best Local Similarity 71.4%; Pred. No. 2e+02;
CC Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 VLVVTSS 7
CC Db 102 VIIVTSS 108
CC |::|||

Search completed: July 16, 2003, 13:13:13
Job time : 5.78049 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:07:18 ; Search time 18.9512 Seconds
(without alignments)
76.107 Million cell updates/sec

Title: US-09-910-582B-4

Perfect score: 29

Sequence: 1 VVLVTSS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	302	10 Q949E0	Q949E0 oryza sativ
2	29	100.0	388	16 Q67179	Q67179 aquifex aeo
3	28	96.6	232	16 Q97ER8	Q97ER8 clostridium
4	28	96.6	265	16 Q9X0W0	Q9X0W0 thermotoga
5	28	96.6	277	17 Q57757	Q57757 pyrococcus
6	28	96.6	277	17 Q9V200	Q9V200 pyrococcus
7	28	96.6	277	17 Q8T2M1	Q8T2M1 pyrococcus
8	28	96.6	341	2 Q9JRR7	Q9JRR7 actinobacil
9	27	93.1	183	5 Q9VUV2	Q9VUV2 drosophila
10	27	93.1	588	10 Q9LT69	Q9LT69 arabidopsis
11	26	89.7	115	2 Q9F3U7	Q9F3U7 pseudomonas
12	26	89.7	173	11 Q9B663	Q9B663 mus musculu
13	26	89.7	177	5 Q9VTC9	Q9VTC9 drosophila
14	26	89.7	198	16 P73908	P73908 synechocyst
15	26	89.7	233	13 Q9PVV9	Q9PVV9 oryzias lat
16	26	89.7	239	10 Q9FHM6	Q9FHM6 arabidopsis

17	26	89.7	242	10	Q9XIG4	Q9XIG4 arabidopsis
18	26	89.7	255	9	Q94MX8	Q94MX8 haemophilus
19	26	89.7	294	16	Q8VYL8	Q8VYL8 anabaena sp
20	26	89.7	302	6	Q9N292	Q9N292 pongo pygma
21	26	89.7	305	2	Q68968	Q68968 synechococc
22	26	89.7	305	13	Q9PVP2	Q9PVP2 oryzias lat
23	26	89.7	313	13	Q90426	Q90426 brachydanio
24	26	89.7	344	10	Q8RWH8	Q8RWH8 arabidopsis
25	26	89.7	349	5	Q16356	Q16356 caenorhabdi
26	26	89.7	351	5	Q968A2	Q968A2 caenorhabdi
27	26	89.7	449	2	Q936J6	Q936J6 rhodospiril
28	26	89.7	454	5	Q45096	Q45096 caenorhabdi
29	26	89.7	559	8	Q9MQ28	Q9MQ28 laqueus rub
30	26	89.7	631	11	Q9ES20	Q9ES20 rattus norv
31	26	89.7	633	11	Q54935	Q54935 cricetus
32	26	89.7	648	11	P97715	P97715 rattus norv
33	26	89.7	727	16	Q8YZG7	Q8YZG7 anabaena sp
34	26	89.7	799	4	Q8VWV8	Q8VWV8 homo sapien
35	26	89.7	799	4	Q9H9B7	Q9H9B7 homo sapien
36	26	89.7	831	10	Q65673	Q65673 arabidopsis
37	26	89.7	860	5	Q9NXX2	Q9NXX2 bombyx mori
38	26	89.7	861	5	Q9NXX3	Q9NXX3 bombyx mori
39	26	89.7	873	13	Q9PUE4	Q9PUE4 brachydanio
40	26	89.7	873	13	Q918E6	Q918E6 fugu rubrip
41	26	89.7	874	11	Q9QZE5	Q9QZE5 mus musculus
42	26	89.7	874	11	Q8R1A7	Q8R1A7 mus musculus
43	26	89.7	879	5	Q9V9W9	Q9V9W9 drosophila
44	26	89.7	879	5	Q9U677	Q9U677 drosophila
45	26	89.7	1331	13	Q90W93	Q90W93 poecilia re

ALIGNMENTS

RESULT 1

Q949E0	ID	Q949E0	PRELIMINARY;	PRT;	302 AA.
AC	Q949E0;				
DT	01-DEC-2001	(TReMBLrel. 19, Created)			
DT	01-DEC-2001	(TReMBLrel. 19, Last sequence update)			
DT	01-MAR-2002	(TReMBLrel. 20, Last annotation update)			
DE	Putative enoyl-CoA hydratase.				
GN	W495ERIPDM.				
OS	Oryza sativa (Rice).				
OC	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;				
OC	Ehrhartoideae; Oryzeae; Oryza.				
OX	NCBI_TaxID=4530;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RP	MEDLINE=21329048; PubMed=11435398;				
RA	Mayer K., Murphy G., Tarchini R., Wambutt R., Voickaert G., Pohl T.,				
RA	Dueterhoeft A., Stiekema W., Entian K.D., Terryn N., Lemcke K.,				
RA	Haase D., Hall C.R., van Dodeweerd A.M., Tingey S.V., Mewes H.W.,				
RA	Bevan M., Bancroft I.;				
RT	"Conservation of microstructure between a sequenced region of the				
RT	genome of rice and multiple segments of the genome of Arabidopsis				
RT	thaliana.";				
RL	Genome Res. 11:1167-1174 (2001).				
DR	EMBL; AJ307662; CAC39052.1;				
DR	InterPro; IPR001753; EnCoA_hydrase.				
DR	Pfam; PF00378; ECH; 1.				
DR	PROSITE; PS00166; ENOYL_COA_HYDRATASE; UNKNOWN_1.				
SQ	SEQUENCE 302 AA; 32201 MW; 90B07A741EEDACBD CRC64;				

Query Match 100.0%; Score 29; DB 10; Length 302;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VVLVTSS.7
Db 130 VVLVTSS 136

Wed Jul 16 13:41:37 2003

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 2
O67179 ID O67179 PRELIMINARY; PRT; 398 AA.
AC O67179;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE HEMOLYSIN
GN HLA OR AQ_1091.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RA "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
DR EMBL; AC000722; AAC07135.1;
DR InterPro; IPR001343; Hemlysin_Ca_bind.
DR InterPro; IPR003355; RTXtoxin_N.
DR Pfam; PF00353; hemolysinCaBind; 4.
DR PRINTS; PR00313; CABNDNGRPT.
DR PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 2.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 398 AA; 40314 MW; 1798DF142CE2BA26 CRC64;

Query Match 100.0%; Score 29; DB 16; Length 398;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
DB 14 VLVLTSS 20
|||||

RESULT 3
O97ER8 ID O97ER8 PRELIMINARY; PRT; 232 AA.
AC O97ER8;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CPSC/CAPB subfamily ATPase.
GN CAC3040.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R.L., Lee H.M., Dubois J., Olu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RA "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AC007801; AAK80980.1;
DR InterPro; IPR000392; NitrogenaseII.
DR Pfam; PF00442; fer4_NifH; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 232 AA; 25475 MW; 6962E2ECA7055EAC CRC64;

Query Match 96.6%; Score 28; DB 16; Length 232;
Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
DB 148 VLVLTSS 154
|||||

RESULT 5
O57757 ID O57757 PRELIMINARY; PRT; 277 AA.
AC O57757;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE 277AA long hypothetical sugar transport membrane protein.
GN PH0023.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RA "Complete sequence and gene organization of the genome of a hyper-
RT

```

```

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
DB 34 VLVLTSS 40
|||||

RESULT 4
O9XOW0 ID O9XOW0 PRELIMINARY; PRT; 265 AA.
AC O9XOW0;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Sugar ABC transporter, permease protein, putative.
GN TM1233.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RL genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
DR EMBL; AC001779; AAD36308.1;
DR TIGR; TM1233;
DR InterPro; IPR000515; BPD_transp.
DR Pfam; PF00528; BPD_transp; 1.
DR PROSITE; PS00402; BPD_TRANS_INN_MEMBR; UNKNOWN_1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 265 AA; 29166 MW; 55A36C188B949B81 CRC64;

Query Match 96.6%; Score 28; DB 16; Length 265;
Best Local Similarity 85.7%; Pred. No. 91;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
DB 148 VLVLTSS 154
|||||

```

RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
 RL DNA Res. 5:55-76(1998).
 DR EMBL; AP000001; BAA29091.1; -.
 DR InterPro: IPR000515; BPD_transp.
 DR Pfam: PF00528; BPD_transp.1.
 DR PROSITE; PS00402; BPD_TRANSF_INN_MEMBR; UNKNOWN_1.
 KW Sugar transport; Complete proteome.
 SQ SEQUENCE 277 AA; 30507 MW; 3F03778650EA8E2A CRC64;

Query Match 96.6%; Score 28; DB 17; Length 277;
 Best Local Similarity 85.7%; Pred. No. 95;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVLTSS 7
 |||:||||
 Db 155 VLVLTSS 161

RESULT 6

Q9V2Q0 PRELIMINARY; PRT; 277 AA.
 ID Q9V2Q0
 AC Q9V2Q0;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Sugar ABC transporter, permease protein.
 GN PAB2335.
 GN PAB2335.
 OS Pyrococcus abyssi.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=29292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ORSAY;
 RA Hellig R.;
 RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure and evolution."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ248283; CAB48948.1; -.
 DR InterPro: IPR000515; BPD_transp.
 DR Pfam; PF00528; BPD_transp.1.
 DR PROSITE; PS00402; BPD_TRANSF_INN_MEMBR; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 277 AA; 30381 MW; EE1DDF00001DF954 CRC64;

Query Match 96.6%; Score 28; DB 17; Length 277;
 Best Local Similarity 85.7%; Pred. No. 95;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVLTSS 7
 |||:||||
 Db 155 VLVLTSS 161

RESULT 7

Q8TZM1 PRELIMINARY; PRT; 277 AA.
 ID Q8TZM1
 AC Q8TZM1;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Sugar ABC transporter, permease protein.
 GN PFI969.
 GN PFI969.
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=2261;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VCL / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the Pyrococcus furiosus genome."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AE010290; AAL82093.1; -.
 KW Complete proteome.
 SQ SEQUENCE 277 AA; 30386 MW; 0D09E02F176EE49D CRC64;

Query Match 96.6%; Score 28; DB 17; Length 277;
 Best Local Similarity 85.7%; Pred. No. 95;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVLTSS 7
 |||:||||
 Db 155 VLVLTSS 161

RESULT 8

Q9JRR7 PRELIMINARY; PRT; 341 AA.
 ID Q9JRR7
 AC Q9JRR7;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE GDP-mannose pyrophosphorylase homolog.
 GN MNB.
 GN MNB.
 OS Actinobacillus actinomycetemcomitans (Haemophilus actinomycetemcomitans).
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Actinobacillus.
 OX NCBI_TaxID=714;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IDH 781;
 RX MEDLINE=20435309; PubMed=10978535;
 RA Nakano Y., Yoshida Y., Suzuki N., Yamashita Y., Koga T.;
 RT "A gene cluster for the synthesis of serotype d-specific polysaccharide antigen in Actinobacillus actinomycetemcomitans."
 RL Biochim. Biophys. Acta 1493:259-263(2000).
 DR EMBL; AB041266; BAA94400.1; -.
 DR InterPro: IPR001538; Man6P_isomerII.
 DR InterPro; IPR001825; NTP_transferase.
 DR Pfam; PF01050; Mannosep_isomer.1.
 DR Pfam; PF00483; NTP_transferase.1.
 SQ SEQUENCE 341 AA; 38833 MW; 6EBD3C9B0B9DFBF1 CRC64;

Query Match 96.6%; Score 28; DB 2; Length 341;
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLVLTSS 7
 :|||||
 Db 99 IVLVTSS 105

RESULT 9

Q9VUV2 PRELIMINARY; PRT; 183 AA.
 ID Q9VUV2
 AC Q9VUV2;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE CG6112 protein.
 GN CG6112.
 GN Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

```

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu K., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
RA de Pablo J.B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Howland K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kaulush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Rao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003529; AAP49571.1; -.
DR FlyBase; FBgn0036542; CG6112.
SQ SEQUENCE 183 AA; 20998 MW; 828BCD3F56D6487C CRC64;

Query Match 93.1%; Score 27; DB 5; Length 183;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
DB 35 IILVTS 41

RESULT 10
Q9LT69 PRELIMINARY; PRT; 588 AA.
AC Q9LT69;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Phosphoglycerate dehydrogenase.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty pl and TAC
RT clones.";

DNA Res. 7:131-135(2000).
DR EMBL; AB025624; BAB02473.1; -.
DR HSSP; P08328; 1PSD.
DR InterPro; IPR002912; ACT.
DR InterPro; IPR002162; D_2hyd_dh.
DR Pfam; PF00389; 2-Hacid_DH; 1.
DR Pfam; PF02826; 2-Hacid_DH_C; 1.
DR Pfam; PF01842; ACT; 1.
DR PROSITE; PS00065; D_2-HYDROXYACID_DH.1; UNKNOWN_1.
DR PROSITE; PS00670; D_2-HYDROXYACID_DH.2; UNKNOWN_1.
DR PROSITE; PS00671; D_2-HYDROXYACID_DH.3; 1.
SQ SEQUENCE 588 AA; 62122 MW; E4C9A35BE86D3AD5 CRC64;

Query Match 93.1%; Score 27; DB 10; Length 588;
Best Local Similarity 71.4%; Pred. No. 3.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
DB 33 IILVTS 39

RESULT 11
Q9F307 PRELIMINARY; PRT; 115 AA.
AC Q9F307;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE MerT protein.
GN MER.
OS Pseudomonas sp. BW13.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=82760;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BW13;
RA Minakhin L.S.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y17830; CAC14713.1; -.
DR InterPro; IPR003457; Transprt_MerT.
DR Pfam; PF02411; MerT; 1.
SQ SEQUENCE 115 AA; 12438 MW; 21C347A580F4C55B CRC64;

Query Match 89.7%; Score 26; DB 2; Length 115;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVLTSS 7
DB 102 LVLVTS 108

RESULT 12
Q9D663 PRELIMINARY; PRT; 173 AA.
AC Q9D663;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 0 day neonate skin cDNA, RIKEN full-length enriched library,
DE clone:4632427M03, full insert sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=SKIN;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
```

RA Salto T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schirni L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Hayashaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK014595; BAB29451.1; -
DR InterPro; IPR002553; Adaptin_N.
DR Pfam; PF01602; Adaptin_N; 1.
SQ SEQUENCE 173 AA; 19306 MW; A63FB7012FDCD74 CRC64;

Query Match 89.7%; Score 26; DB 11; Length 173;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVVTSS 7
|:|||||
DB 30 VIIVTSS 36

RESULT 13

Q9VTC9 ID Q9VTC9 PRELIMINARY; PRT; 177 AA.

AC Q9VTC9

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)

DE CG6391 protein.

GN CG6391

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton R.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbini K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003546; AAF50123.1; -
DR FlyBase; FBgn0036111; CG6391.
DR InterPro; IPR000086; NUDIX_hydrolase.
DR Pfam; PF00293; NUDIX; 1.
DR PROSITE; PS00893; NUDIX; 1.
SQ SEQUENCE 177 AA; 19879 MW; CE0AD5232B7C04B3 CRC64;

Query Match 89.7%; Score 26; DB 5; Length 177;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVVTSS 7
|:|||||
DB 34 VLVVTSS 40

RESULT 14

P73908 ID P73908 PRELIMINARY; PRT; 198 AA.

AC P73908

DT 01-FEB-1997 (TrEMBLrel. 02, Created)

DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Cytochrome C oxidase subunit III.

GN CTAE OR SLR2083.

OS Synechocystis sp. (strain PCC 6803).

OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.

OX NCBI_TaxID=1148;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97061201; PubMed=8905231;

RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirose M., Sugita M., Sugita M., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; D90910; BAA17972.1; -
DR InterPro; IPR002982; CytC_oxdase_III.
DR Pfam; PF00510; COX3; 1.
DR ProDom; PD000382; CytC_oxdase_III; 1.
DR PROSITE; PS02553; COX3; 1.
KW Complete proteome.
SQ SEQUENCE 198 AA; 22177 MW; E0B5F6FA66CADB5B CRC64;

Query Match 89.7%; Score 26; DB 16; Length 198;
Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLVVTSS 7
|:|||||
DB 71 MVLVTSS 77

RESULT 15

Q9PVV9

ID Q9PVV9 PRELIMINARY; PRT; 233 AA.
 AC Q9PVV9;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
 DE Olfactory receptor 4 (Fragment).
 GN MFOR4.
 OS Oryzias latipes (Medaka fish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
 OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
 OX NCBI_TaxID=8090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NGY;
 RX MEDLINE=99250259; PubMed=10231578;
 RA Sun H., Kondo R., Shima A., Naruse K., Hori H., Chigusa S.I.;
 RT "Evolutionary analysis of putative olfactory receptor genes of medaka
 fish, Oryzias latipes.";
 RL Gene 231:137-145(1999).
 DR EMBL; AB022645; BAA84278.1; .
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 FT NON_TER 1 1
 FT NON_TER 233 233
 SQ SEQUENCE 233 AA; 26325 MW; 7FIAC6729550089DE CRC64;

 Query Match 89.7%; Score 26; DB 13; Length 233;
 Best Local Similarity 85.7%; Pred. No. 2.4e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 1 VVLVTSS 7
 Db 178 VVLLTSS 184

Search completed: July 16, 2003, 13:15:13
 Job time : 20.9512 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:04:13 ; Search time 27.3171 Seconds
(without alignments)
39.023 Million cell updates/sec

Title: US-09-910-582B-9
Perfect score: 51
Sequence: 1 CLHRGNSC 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	8	22	Peptide which sele
2	51	100.0	8	22	Heart homing pepti
3	38	74.5	193	22	Propionibacterium
4	36	70.6	43	21	Human secreted pro
5	36	70.6	51	23	Human ORF3867 prot
6	36	70.6	55	22	Propionibacterium
7	36	70.6	93	22	Propionibacterium
8	36	70.6	99	19	Partial mouse immu
9	36	70.6	155	21	Human secreted pro
10	36	70.6	224	23	Murine protein iso

11	36	70.6	329	23	AAU80023	Human K-ras gene p
12	36	70.6	4636	22	AAE11937	Human CG168 (or C5
13	35.5	69.6	879	18	AAW31264	Human DRADA mutat
14	35.5	69.6	1226	16	AAW79346	Human 140 kD DRADA
15	35.5	69.6	1226	18	AAW31232	Human double stran
16	35.5	69.6	1226	19	AAW54962	Human double-stran
17	35	68.6	98	22	AAW4681	Human Immune/haema
18	35	68.6	130	20	AAV11509	Human 5' EST secre
19	35	68.6	133	22	ABG10765	Novel human diagno
20	35	68.6	173	22	AAU54669	Propionibacterium
21	35	68.6	256	22	AAU60445	Propionibacterium
22	35	68.6	532	22	AAU53526	Propionibacterium
23	35	68.6	794	22	ABG08768	Novel human diagno
24	35	68.6	920	22	AAE09955	Methylomonas 16a n
25	35	68.6	920	23	ABG61576	High growth methan
26	35	68.6	967	22	ABG16000	Novel human diagno
27	35	68.6	1447	22	AAE86975	D. melanogaster la
28	35	68.6	1905	22	ABE59243	Drosophila melanog
29	35	68.6	1905	22	AAU38925	Drosophila G-prote
30	35	68.6	3070	23	AAO17359	Human laminin M ch
31	35	68.6	3084	21	AAE19796	Mouse laminin 2 ma
32	35	68.6	3088	21	AAE19794	Human laminin 2 ma
33	35	68.6	3089	21	AAE19792	Human laminin 2 ma
34	35	68.6	3106	21	AAE19795	Human laminin 2 al
35	35	68.6	3110	16	AAE71730	Merosin major subu
36	35	68.6	3110	20	AAV15460	Human laminin alph
37	35	68.6	3110	21	AAE19791	Human laminin 2 al
38	35	68.6	3110	21	AAE19793	Human laminin 2 al
39	35	68.6	3110	23	AAU84345	Protein LAM2 diff
40	35	68.6	3150	22	ABG20414	Novel human diagno
41	34	66.7	14	21	AAE34081	Human secreted pro
42	34	66.7	34	19	AAW79024	Rieske motif in PO
43	34	66.7	44	21	AAE34515	Human secreted pro
44	34	66.7	44	22	AAU01679	Gene 23 human secr
45	34	66.7	69	22	AAU44905	Propionibacterium

ALIGNMENTS

RESULT 1

AAE30898

ID AAB30898 standard; peptide; 8 AA.

XX AAB30898;

XX AC

XX DT 02-APR-2001 (first entry)

XX DE

XX DE Peptide which selectively binds to normal cardiac endothelium.

XX DE Cardiac endothelium; angiogenic factor; vascular endothelium;

XX KW Cardiac endothelium; angiogenic factor; vascular endothelium;

XX KW peripheral vascular disease; cardiovascular disease; angiogenesis;

XX KW cardiac neovascularisation.

XX OS

XX OS Unidentified.

XX PN

XX PN WO200075329-A1.

XX PD

XX PD 14-DEC-2000.

XX PF

XX PF 31-MAY-2000; 2000WO-US14988.

XX PR

XX PR 07-JUN-1999; 99US-0327045.

XX PA

XX PA (EDWA-) EDWARDS LIFESCIENCES CORP.

XX PA (BAXT) BAXTER AG.

XX PI

XX PI Levine AJ, Mitterer A, Falkner F, Scheiflinger F, Dorner F;

XX DR

XX DR WPI; 2001-091212/10.

XX PT

XX PT New chimeric molecules having an angiogenic factor linked to a

XX PT targeting molecule that binds to a vascular endothelium, useful for

PT increasing cardiac neovascularisation, or treating peripheral vascular
 PT and cardiovascular diseases -
 XX Disclosure; Page 27; 67pp; English.
 PS
 XX AAB30895-99 represent targeting molecules, which are used to produce the
 CC chimeric molecules of the invention. AAB30895-98 selectively bind to
 CC normal cardiac endothelium. The specification describes a chimeric
 CC molecule comprising an angiogenic factor linked to a targeting molecule
 CC that specifically binds to a vascular endothelium. The chimeric
 CC molecules are useful for treatment of peripheral vascular or
 CC cardiovascular diseases. Specifically, they are useful for inducing or
 CC inhibiting angiogenesis, for increasing cardiac neovascularisation in
 CC ischemic tissue in the peripheral vascular system.
 XX
 XX Sequence 8 AA;
 SQ
 Query Match 100.0%; Score 51; DB 22; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CLHRGNSC 8
 Db | | | | | | | |
 1 CLHRGNSC 8
 RESULT 2
 AAB59303
 ID AAB59303 standard; peptide; 8 AA.
 XX
 AC AAB59303;
 XX
 XX 21-MAR-2001 (first entry)
 DT
 DE Heart homing peptide SEQ ID NO: 9.
 XX
 XX Heart homing peptide; cardiovascular disease; ischaemic disease;
 KW gene therapy.
 KW
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Disulfide-bond 1..8
 FT /note= "disulfide bond cyclises the peptide"
 FT
 XX WO200075174-A1.
 PN
 XX 14-DEC-2000.
 PD
 XX 31-MAY-2000; 2000WO-US15088.
 PF
 XX 07-JUN-1999; 99US-0326718.
 PR
 XX (BURN-) BURNHAM INST.
 PA
 XX Ruoslahti E, Mackenna DA;
 PI
 XX WPI: 2001-071059/08.
 DR
 XX Novel heart homing peptide that selectively homes to normal ischaemic
 XX and cardiac tissue useful for targeting ischaemic tissues for treating
 PT ischaemic and cardiovascular diseases such as atherosclerosis and
 PT restenosis -
 PT
 XX Claim 2; Page 55; 70pp; English.
 PS
 XX The present invention provides a number of heart homing peptides which
 CC selectively home to cardiac tissue. These can be used in the treatment of
 CC cardiovascular and ischaemic diseases, such as atherosclerosis, CC
 CC thrombosis, restenosis, vasculitis, atherosclerotic aneurysms, myocardial
 CC hypertrophy, congenital heart diseases, ischaemic heart disease and
 CC anginas, acquired valvular/endocardial diseases, primary myocardial
 CC diseases, cardiac tumours and arrhythmias.

XX Sequence 8 AA;
 SQ
 Query Match 100.0%; Score 51; DB 22; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CLHRGNSC 8
 Db | | | | | | | |
 1 CLHRGNSC 8
 RESULT 3
 AAU40780
 ID AAU40780 standard; Protein: 193 AA.
 XX
 AC AAU40780;
 XX
 XX 13-FEB-2002 (first entry)
 DT
 DE Propionibacterium acnes immunogenic protein #1676.
 XX
 XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 XX WO200181581-A2.
 PN
 XX 01-NOV-2001.
 PD
 XX 20-APR-2001; 2001WO-US12865.
 PF
 XX 21-APR-2000; 2000US-199047P.
 PR
 XX 02-JUN-2000; 2000US-208841P.
 PR
 XX 07-JUL-2000; 2000US-216747P.
 PR
 XX (CORI-) CORIXA CORP.
 PA
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 PI
 XX WPI: 2001-616774/71.
 DR
 XX N-PSDB; AAS559513.
 DR
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 PT
 XX Example 1; SEQ ID NO 1975; 1069pp; English.
 PS
 XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 193 AA;
 Query Match 74.5%; Score 38; DB 22; Length 193;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HRGNSC 8
 |||||
 Db 87 HRGNSC 92

RESULT 4
 AAB32465
 ID AAB32465 standard; Protein; 43 AA.
 AC AAB32465;
 DT 16-JAN-2001 (first entry)
 DE Human secreted protein sequence encoded by gene 22 SEQ ID NO:151.
 KW Human; secreted protein; cytostatic; immunostimulant; antiproliferative;
 KW cardiant; antiarrhythmic; antiviral; antibacterial; antifungal; cancer;
 KW antiparasitic; neuroprotective; nontropic; antiinflammatory; anti-HIV;
 KW antiangiogenic; antiarteriosclerotic; diagnosis; immune disorder; AIDS;
 KW autoimmune disease; haematopoietic cell disorder; blood protein disorder;
 KW agammaglobulinaemia; hyperproliferative disease; Gaucher's disease;
 KW cardiovascular disease; congenital heart defect; pulmonary atresia;
 KW arrhythmia; ischaemia; angiogenesis related disorder; Crohn's disease;
 KW atherosclerosis; neurological disease; Alzheimer's disease;
 KW Huntington's; infectious disease; cat-scratch disease.
 OS Homo sapiens.
 XX
 XX
 PN WO200047602-A1.
 XX
 XX 17-AUG-2000.
 XX
 XX 08-FEB-2000; 2000WO-US03062.
 XX
 XX 10-FEB-1999; 99US-0119468.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Ruben SM, Ebner R, Young PE, Ni J, Soppet DR, Moore PA;
 PI Shi Y, Lafleur DW, Olsen HS, Florence KA, Komatsoulis G;
 XX
 XX WPI; 2000-543578/49.
 XX
 XX New human nucleic acids encoding secreted proteins, useful in the
 XX treatment, prevention or diagnosis of immune disorders (e.g. autoimmune
 XX diseases), blood protein disorders and hyperproliferative diseases
 XX (e.g. Gaucher's disease) -
 XX
 XX Disclosure; Page 56; 488pp; English.

XX CC The polynucleotide sequences given in AAC55190 to AAC55235 encode the
 XX human secreted proteins given in AAB32371 to AAB32484. Human secreted
 XX proteins have activities based on the tissues and cells the genes are
 XX expressed in. Examples of activities include: cytostatic;
 XX immunostimulant; antiproliferative; cardiant; antiarrhythmic; antiviral;
 XX antibacterial; antifungal; antiparasitic; neuroprotective; nontropic;
 XX antiinflammatory; antiangiogenic; anti-HIV; and antiarteriosclerotic. The
 XX polynucleotides and polypeptides, or their agonists and antagonists, can
 XX be used for treating, preventing or diagnosing immune disorders (e.g.
 XX cancer, autoimmune diseases), disorders of haematopoietic cells, blood
 XX protein disorders (e.g. agammaglobulinaemia), hyperproliferative
 XX diseases (e.g. Gaucher's disease), cardiovascular disorders (e.g.
 XX congenital heart defects, pulmonary atresia, arrhythmias, ischaemia).
 XX angiogenesis related disorders (e.g. Crohn's disease, atherosclerosis),
 XX neurological diseases (e.g. Alzheimer's disease, Huntington's chorea),
 XX infectious diseases (e.g. AIDS, cat-scratch disease and other bacterial,

CC viral, parasitic or fungal diseases). AAC55181 to AAC55189 and AAB32370
 CC represent sequences used in the exemplification of the present
 CC invention.
 XX SQ Sequence 43 AA;
 Query Match 70.6%; Score 36; DB 21; Length 43;
 Best Local Similarity 71.4%; Pred. No. 33;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LHRGNSC 8
 :||||:
 Db 23 VHRGNTC 29

RESULT 5
 ABP34894
 ID ABP34894 standard; Protein; 51 AA.
 XX
 AC ABP34894;
 DT 08-JUL-2002 (first entry)
 DE Human ORF3867 protein, SEQ ID NO:7734.
 KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 KW vasotropic; antipsoriatic; antidiabetic; cytostatic; nontropic;
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.
 XX
 OS Homo sapiens.
 XX
 XX WO200190366-A2.
 XX
 XX 29-NOV-2001.
 XX
 XX 24-MAY-2001; 2001WO-US17076.
 XX
 XX 24-MAY-2000; 2000US-206690P.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 XX Leach MD, Shinkets RA;
 PI
 XX WPI; 2002-106200/14.
 DR
 DR N-PSDB; ABN78920.
 XX
 XX Novel human polypeptides and polynucleotides useful for diagnosing,
 XX preventing and treating cardiovascular disease, neurodegenerative,
 XX hyperproliferative disorders and disorders related to organ
 XX transplantation -
 XX
 XX Claim 10; Page 2177-2178; 2508pp; English.

XX CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 XX designated ORF (Open reading frame) 1-4534, and sequences ABN75054-
 XX ABN79587 represent cDNAs encoding them. The invention also encompasses
 XX polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 XX referred to as ORFX) proteins, polynucleotides at least 85% identical to
 XX the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 XX polynucleotides, the recombinant production of ORFX proteins, antibodies
 XX specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 XX polypeptides, methods of screening for modulators of ORFX expression or
 XX activity, and methods of screening individuals for a predisposition to an

ORFX-associated disorder. The ORFX proteins of the invention have a wide range of biological activities, such as cytokine, cell proliferation, cell differentiation, immune modulation, haematopoiesis regulation, tissue growth, angiogenesis, activin or inhibin activity, chemotactic/chemokinetic activity, haemostatic activity, thrombolytic activity, receptor/ligand, antiinflammatory activity, tumour inhibition activity, and antifertility activity, and may also be involved in the determination of bodily characteristics, fertility and behaviour. ORFX proteins, nucleic acids and antibodies may be used in the treatment of cancers, neuroproliferative disorders such as epilepsy and Alzheimer's disease, cardiovascular diseases, immune system disorders, disorders related to organ transplantation, disorders of tissue growth and regeneration, diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester storage disease, and infectious diseases caused by viral, bacterial, fungal and other pathogens. ORFX nucleic acids may also be used as a source of primers and probes, in the detection of ORFX genomic sequences or transcripts, in the identification and cloning of homologous sequences, in genetic diagnosis, and in forensic biology. The ORFX nucleic acids may additionally be used to produce transgenic animals which may be useful for studying the function and/or activity of ORFX protein, and in drug screening. The ORFX proteins may also be used as immunogens to generate specific antibodies, which are useful in the diagnosis, treatment and monitoring of ORFX-associated diseases.

Query Match 70.6%; Score 36; DB 23; Length 51;
Best Local Similarity 62.5%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
Db 24 CVHRGEFC 31

RESULT 6
AAU52817
ID AAU52817 standard; Protein; 55 AA.

XX AAU52817;
XX 27-FEB-2002 (first entry)
XX Propionibacterium acnes immunogenic protein #13713.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.
XX WO200181581-A2.

XX 01-NOV-2001.
XX 20-APR-2001; 2001WO-US12865.
XX 21-APR-2000; 2000US-199047P.
XX 02-JUN-2000; 2000US-208841P.
XX 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.
XX N-PSDB; AAS59557.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for

treating acne vulgaris -
Example 1; SEQ ID No 14012; 1069pp; English.

Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic polypeptides. The proteins and their associated DNA sequences are used in the treatment, prevention and diagnosis of medical conditions caused by P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to downregulate expression and activity of P. acnes polypeptides and therefore treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence, for example, by enzyme linked immunosorbent assay (ELISA).
Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 55 AA;

Query Match 70.6%; Score 36; DB 22; Length 55;
Best Local Similarity 75.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
Db 39 CLQRGXSC 46

RESULT 7
AAU41229
ID AAU41229 standard; Protein; 93 AA.

XX AAU41229;
XX 13-FEB-2002 (first entry)
XX Propionibacterium acnes immunogenic protein #2125.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.
XX WO200181581-A2.

XX 01-NOV-2001.
XX 20-APR-2001; 2001WO-US12865.
XX 21-APR-2000; 2000US-199047P.
XX 02-JUN-2000; 2000US-208841P.
XX 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.
XX N-PSDB; AAS59514.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for

PT treating acne vulgaris -

PS Example 1; SEQ ID NO 2424; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic polypeptides. The proteins and their associated DNA sequences are used in the treatment, prevention and diagnosis of medical conditions caused by P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to downregulate expression and activity of P. acnes polypeptides and therefore treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence, for example, by enzyme linked immunosorbent assay (ELISA).

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 93 AA;

Query Match 70.6%; Score 36; DB 22; Length 93;

Best Local Similarity 85.7%; Pred. No. 66;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CLHRGNS 7

Db 34 CLHRGQS 40

RESULT 8

AAW62773

ID AAW62773 standard; Protein; 99 AA.

AC AAW62773;

XX 23-SEP-1998 (first entry)

DE Partial mouse immunoglobulin receptor designated FDF03.

XX Human; type I transmembrane protein; immunoglobulin-like domain;

KW FDF03; activated monocyte; YE01; KTE03; control; development;

KW differentiation; mammalian immune system; treatment; cancerous condition;

KW degenerative condition; autoimmune response; transplantation rejection;

KW graft versus host disease; inflammatory condition; detection; diagnosis;

KW drug screening.

XX Mus sp.

OS WO9824906-A2.

PN 11-JUN-1998.

XX 05-DEC-1997; 97WO-US21101.

XX 21-MAR-1997; 97US-0041279.

PR 06-DEC-1996; 96US-0032252.

PR 09-DEC-1996; 96US-0762187.

PR 16-DEC-1996; 96US-0033181.

XX (SCHE) SCHERING CORP.

XX Adema GJ, Gorman DM, Lanier LL, McClanahan TK, Meygaard L;

PI Phillips JH, Zurawski G, Zurawski SM;

XX WPI; 1998-333325/29.

DR N-PSDB; AAV38988.

XX

PT New isolated activated monocyte cell gene(s) - used to develop

PT products for treating e.g. cancer, degenerative conditions,

PT autoimmune responses, transplant rejection or inflammatory

PT conditions

XX Claim 1; Pages 62-63; 104pp; English.

XX The present sequence represents a partial rodent protein which is a type I transmembrane protein comprising an extracellular portion characterised by immunoglobulin-like domains, indicating that the protein, FDF03, is a receptor member of the immunoglobulin superfamily. The FDF03 gene is found in activated monocytes. The specification also describes other proteins encoded by activated monocytes, which are designated YE01 and KTE03. The genes function in controlling development, differentiation, and/or physiology of the mammalian immune system. The products can be used for treating abnormal proliferation, regeneration, degeneration or atrophy. They can be used for treating e.g. cancerous conditions, degenerative conditions, autoimmune responses, transplantation rejection, graft versus host disease, or inflammatory conditions. The products can also be used for detection, diagnosis and drug screening.

XX SQ Sequence 99 AA;

Query Match 70.6%; Score 36; DB 19; Length 99;

Best Local Similarity 85.7%; Pred. No. 70;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CLHRGNS 7

Db 12 CLHRGNS 18

RESULT 9

AAB32463

ID AAB32463 standard; Protein; 155 AA.

XX AAB32463;

XX 16-JAN-2001 (first entry)

DE Human secreted protein sequence encoded by gene 22 SEQ ID NO:149.

XX Human; secreted protein; cytostatic; immunostimulant; antiproliferative;

KW cardiant; antiarrhythmic; antiviral; antibacterial; antifungal; cancer;

KW antiparasitic; neuroprotective; neutropic; antiinflammatory; anti-HIV;

KW antiangiogenic; antiarteriosclerotic; diagnosis; immune disorder; AIDS;

KW autoimmune disease; haematopoietic cell disorder; blood protein disorder;

KW agammaglobulinaemia; hyperproliferative disease; Gaucher's disease;

KW cardiovascular disorder; congenital heart defect; pulmonary atresia;

KW arrhythmia; ischaemia; angiogenesis related disorder; Crohn's disease;

KW atherosclerosis; neurological disease; Alzheimer's disease;

KW Huntington's; infectious disease; cat-scratch disease.

XX Homo sapiens.

OS WO200047602-A1.

PN 17-AUG-2000.

XX 08-FEB-2000; 2000WO-US03062.

XX 10-FEB-1999; 99US-0119468.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Ebner R, Young PE, Ni J, Soppet DR, Moore PA;

PI Shi Y, Lafleur DW, Olsen HS, Florence KA, Komatsoulis G;

XX WPI; 2000-543578/49.

XX New human nucleic acids encoding secreted proteins, useful in the treatment, prevention or diagnosis of immune disorders (e.g. autoimmune

PT diseases), blood protein disorders and hyperproliferative diseases
 PR (e.g. Gaucher's disease) -
 XX
 PS Disclosure; Page 55; 488pp; English.
 XX
 PS The polynucleotide sequences given in AAC55190 to AAC55235 encode the
 CC human secreted proteins given in AAB32371 to AAB32484. Human secreted
 CC proteins have activities based on the tissues and cells the genes are
 CC expressed in. Examples of activities include: cytostatic;
 CC immunostimulant; antiproliferative; cardiant; neuroprotective; antitumor;
 CC antibacterial; antifungal; antiparasitic; anti-HIV; and antiarteriosclerotic. The
 CC polynucleotides and polypeptides, or their agonists and antagonists, can
 CC be used for treating, preventing or diagnosing immune disorders (e.g.
 CC cancer, autoimmune diseases), disorders of haematopoietic cells, blood
 CC protein disorders (e.g. agammaglobulinemia), hyperproliferative
 CC diseases (e.g. Gaucher's disease), cardiovascular disorders (e.g.
 CC congenital heart defects, pulmonary atresia, arrhythmias, ischaemia),
 CC angiogenesis related disorders (e.g. Crohn's disease, atherosclerosis),
 CC neurological diseases (e.g. Alzheimer's disease, Huntington's chorea),
 CC infectious diseases (e.g. AIDS, cat-scratch disease and other bacterial,
 CC viral, parasitic or fungal diseases). AAC55181 to AAC55189 and AAB32370
 CC represent sequences used in the exemplification of the present
 CC invention.
 XX
 SQ Sequence 155 AA;

Query Match 70.6%; Score 36; DB 21; Length 155;
 Best Local Similarity 71.4%; Pred. No. 1.1e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 2 LHRGNSC 8
 :|||:|
 Db 66 VHRGNTC 72

RESULT 10
 ABB72387
 ID ABB72387 standard; Protein: 224 AA.
 XX
 AC ABB72387;
 XX
 DT 04-APR-2002 (first entry)
 XX
 DE Murine protein isolated from skin cells SEQ ID NO: 711.
 XX
 KW Human; rat; mouse; skin cell; skin wound; cancer; growth defect;
 KW developmental defect; inflammatory disease; dermatological; vulnary;
 KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.
 XX
 OS Mus sp.
 XX
 PN WO200190357-Al.
 XX
 PD 29-NOV-2001.
 XX
 PF 24-MAY-2001; 2001WO-NZ00099.
 XX
 PR 24-MAY-2000; 2000US-206650P.
 PR 25-JUL-2000; 2000US-221232P.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX
 PI Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG, Kumble KD;
 XX
 DR WPI; 2002-122020/16.
 DR N-PSDB; ABL35077.
 XX

New polynucleotides and polypeptides encoded by the polynucleotides
 PT isolated from skin cells, useful for treating skin wounds, cancers,
 PT growth and developmental defects, inflammatory diseases, or for
 PT modulating immune responses
 XX

PS Claim 4; Page 454; 466pp; English.
 XX
 CC The present invention provides the protein and coding sequences of cDNAs
 CC isolated from human, murine and rat skin cell libraries. The sequences
 CC can be used in the development of therapeutic agents useful in the
 CC treatment of skin diseases, including skin wounds, cancer, growth
 CC defects, developmental defects and inflammatory diseases. The proteins
 CC have important roles in the induction of hair growth, cell proliferation
 CC and cell-cell interaction in maintaining tissue integrity, in wound
 CC healing and in modulating immune responses. The present sequence is a
 CC polypeptide of the invention.
 XX
 SQ Sequence 224 AA;

Query Match 70.6%; Score 36; DB 23; Length 224;
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CLHRGNS 7
 :|||:|
 Db 25 CLHRGNS 31

RESULT 11
 AAU80023
 ID AAU80023 standard; Protein: 329 AA.
 XX
 AC AAU80023;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human K-ras gene protein 36.
 XX
 KW Human; K-ras gene protein 36; malignant tumour; haemopathy; HIV;
 KW human immunodeficiency virus infection; immunological disease;
 KW inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200183538-Al.
 XX
 PD 08-NOV-2001.
 XX
 PF 28-APR-2001; 2001WO-CN00643.
 XX
 PR 29-APR-2000; 2000CN-0115526.
 XX
 PA (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.
 XX
 PI Mao Y, Xie Y;
 XX
 WPI; 2002-062107/08.
 DR N-PSDB; ABK48081.
 XX
 PT Human K-ras gene protein 36 and encoded polynucleotide, applicable in
 PT diagnosis and treatment of malignant tumour, haemopathy, HIV infection,
 PT immunological diseases and various inflammations
 XX
 PS Claim 1; Page 31-32; 40pp; Chinese.
 XX
 CC The invention relates to an isolated polypeptide of human K-ras gene
 CC protein 36. The polypeptide and encoded polynucleotide are applicable in
 CC diagnosis and treatment of malignant tumour, haemopathy, human
 CC immunodeficiency virus (HIV) infection, immunological diseases and
 CC various inflammations. The present sequence represents the amino acid
 CC sequence of human K-ras gene protein 36.
 XX
 SQ Sequence 329 AA;

Query Match 70.6%; Score 36; DB 23; Length 329;
 Best Local Similarity 71.4%; Pred. No. 2.1e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy      2 LHRGNSC 8
      :|||:|
Db     240 VHRGNTC 246
```

RESULT 12
AAE11937

AAE11937
ID AAE11937 standard; Protein; 4636 AA.

AC AAE11937;

DT 18-DEC-20

DE Human CG168 (or C595) rece

Human; apolipoprotein; lipase; lipoprotein

cardiovascular disease; lipid metabolism; myocardial infarction; cerebral ischaemia; arterial thrombosis; thrombolytic; antilipaeamic; coronary artery thrombosis; cerebral artery thrombosis; stroke; intracardiac thrombosis; gene therapy; cardiovascular; vasodilator; neuroprotectant; cerebroprotective.

OS Homo sapiens.

PN WO200179446-A2.

25-OCT-2001.

16-APR-2001: 2001WO-US12529.

PR 14-APR-2000; 2000US-197137P.

PR 03-AUG-2000: 2000US-0631451
Z0-JUN-2000; 2000US-0398042

PR 22-SEP-2000; 2000US-0667298.
PR 17-NOV-2000; 2000US-0714936.

PA (HYSE-) HYSEO INC.

AA Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
PI Wang D;

WPI: 2001-611724/70.

DR N-PSDB; AAD11937.

PT Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein receptor polypeptides, useful for preventing diagnosing and treating lipid metabolism disorders, thrombosis and cardiovascular diseases -

PS Claim 10: Fig 5: 266pp: English.

The invention relates to polynucleotides encoding proteins CG122, CG179, CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins involved in lipid metabolism and cardiovascular disease such as human apolipoproteins, lipases and lipoprotein receptor proteins. These DNA and protein sequences are useful for treating or preventing disorders associated with apolipoproteins, lipases and lipoprotein receptor (ALLR) expression and for treating lipid metabolism, cardiovascular diseases and thrombosis. Antibodies against these proteins are useful for determining the presence of or predisposition to a disease associated with altered levels of these sequences. ALLR polypeptides are also useful for identifying agents (agonists and antagonists) that bind to them and cells expressing ALLR proteins are useful for identifying a therapeutic agent for use in treatment of a pathology related to aberrant expression or physiological interactions of this polypeptide. Vectors comprising these DNA and protein sequences are also useful for producing ALLR proteins. The nucleic acids and polypeptides of the invention are also useful for the treatment of occlusive cardiovascular diseases, myocardial infarction, cerebral ischaemia, angina, arterial thrombosis, coronary artery thrombosis and cerebral artery thrombosis or intracardiac thrombosis and stroke. The nucleotides of the invention are used in gene therapy. The present sequence is human CG168 (or c595) receptor protein.

DR WPI: 1997-511927/47.
 DR N-PSDB; AAR79346.
 XX
 PT Antibodies to human double-stranded RNA adenosine deaminase - useful
 PT for diagnosis of neurological or CNS disorders, and for drug
 PT screening
 XX
 PS Example 5; Page -: 48pp; English.
 XX
 CC This sequence represents a recombinant mutant pVLDRAD140-delta which
 CC has been constructed from the human double-stranded RNA adenosine
 CC deaminase DRADA (see AAW31232) and lacks the C-terminal 346 amino acids.
 CC This construct is used in expression studies of the protein in Spodoptera
 CC frugiperda. Full length DRADA deaminates multiple adenosine residues to
 CC inosines by a hydrolytic deamination reaction creating Inosine-Uracil
 CC (i-U) mismatched base pairs in double stranded RNA (dsRNA). The
 CC accumulation of extensive mismatched i-U base pairs in the dsRNA causes
 CC unwinding of the RNA double helix. As the enzyme introduces changes to
 CC the sequence of its substrate RNA it is thought to be involved in the
 CC RNA editing process. This protein also contains regions which are
 CC N-terminal truncated versions of DRADA which are biologically active with
 CC similar activity to the full length protein and three double stranded RNA
 CC (dsRNA) binding motifs, DRBM1, DRBM2 and DRBM3. Antibodies derived from
 CC this protein may be used for diagnosis of conditions associated with
 CC abnormal, excessive or deficient expression of DRADA e.g. neurological or
 CC disorders such as Alzheimer's disease, Huntington's disease, SSPE
 CC (subacute sclerosing panencephalitis), measles inclusion body
 CC encephalitis and for drug screening.
 CC Note: This sequence does not appear in the specification, it has been
 CC made by modifying the novel human DRADA sequence found in Figure 1.
 XX
 XX Sequence 879 AA;
 SQ
 Query Match 69.6%; Score 35.5; DB 18; Length 879;
 Best Local Similarity 66.7%; Pred. No. 6.3e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 1;
 QY 1 CLHR-GNSC 8
 DB 622 CMHKLNSC 630
 I: I: I I I I I
 RESULT 14
 AAR79346
 ID AAR79346 standard; Protein; 1226 AA.
 XX
 AC AAR79346;
 XX
 DT 22-JUL-1996 (first entry)
 XX
 DE Human 140 kD DRADA protein.
 XX
 KW DRADA; double stranded ribonucleic acid adenosine deaminase;
 KW neurodegenerative disorder; Alzheimer's disease; analogue;
 KW diagnosis; Huntington's disease.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 200..250
 FT /note= "internal repeat region"
 FT Region 445..449
 FT /note= "partial RNP core consensus sequence"
 FT Region 169..185
 FT /note= "bipartite nuclear localisation signal"
 FT Protein 404..1226
 FT /note= "93 kD DRADA protein"
 FT Protein 440..1226
 FT /note= "88 kD DRADA protein"
 FT Domain 797..1226
 FT /label= catalytic_domain
 FT Binding-site 501..573
 FT /label= DRBM1

FT Binding-site /note= "double stranded RNA binding motif 1"
 FT 613..684
 FT /label= DRBM2
 FT /note= "double stranded RNA binding motif 2"
 FT 725..796
 FT /label= DRBM3
 FT /note= "double stranded RNA binding motif 3"
 XX
 PN WO9522604-A1.
 XX
 PD 24-AUG-1995.
 XX
 PF 16-FEB-1995; 95WO-US02275.
 XX
 PR 25-JUL-1994; 94US-0280443.
 PR 17-FEB-1994; 94US-0197794.
 XX
 PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
 XX
 PI Nishikura K;
 XX
 DR WPI: 1995-302713/39.
 DR N-PSDB; AAR97228.
 XX
 PT Human double stranded ribonucleotide acid adenosine deaminase
 PT enzyme, DRADA - useful in treating neuro-degenerative disorder(s)
 PT e.g. Alzheimer's disease, etc.
 PT
 PS Claim 1; Page 47-52; 98pp; English.
 XX
 CC AAR79346 represents the human double stranded ribonucleic acid
 CC adenosine deaminase enzyme (DRADA). The DRADA protein or
 CC fragments of the protein, polynucleotide sequence and DRADA
 CC antibodies are useful in the diagnosis of certain neurological
 CC or central nervous system disorders e.g. Alzheimer's disease,
 CC Huntington's disease, subacute sclerosing panencephalitis (SSPE),
 CC measles inclusion body encephalitis, strokes, etc. The DRADA protein
 CC or protein fragments may be used to correct the malfunctioning of
 CC defects in glutamate-gated ion channels which result in Alzheimer's
 CC disease, seizures or strokes.
 XX
 SQ Sequence 1226 AA;
 Query Match 69.6%; Score 35.5; DB 16; Length 1226;
 Best Local Similarity 66.7%; Pred. No. 8.5e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 1;
 QY 1 CLHR-GNSC 8
 DB 622 CMHKLNSC 630
 I: I: I I I I I
 RESULT 15
 AAW31232
 ID AAW31232 standard; Protein; 1226 AA.
 XX
 AC AAW31232;
 XX
 DT 26-FEB-1998 (first entry)
 XX
 DE Human double stranded adenosine deaminase.
 XX
 KW Double stranded adenosine deaminase; DRADA; RNA editing; dsRNA;
 KW double stranded RNA binding motif; hydrolytic deamination.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1..1226
 FT /note= "140 kD DRADA protein"
 FT Region 169..185
 FT /note= "Bipartite nuclear localisation signal
 FT comprising two basic residues followed by

FT ten flanking residues and a basic cluster"
FT 200...250
FT /label= internal_repeat
FT /note= "This repeat seems to be unique to humans"
FT 404...1226
FT /note= "N-truncated 93 kD DRADA protein which
FT retains dsRNA deaminase activity"
FT 440...1226
FT /note= "N-truncated 88 kD DRADA protein which
FT retains dsRNA deaminase activity"
FT 445...449
FT /label= RNP2-like_core_consensus
FT /note= "This short RNP2-like stretch may
FT participate in destabilising A-U base pairs
FT and in creating a local single stranded RNA
FT region before adenosine deaminase"
FT 502...573
FT /label= DRBM1
FT /note= "dsRNA binding motif"
FT 613...684
FT /label= DRBM2
FT /note= "dsRNA binding motif"
FT 725...796
FT /label= DRBM3
FT /note= "dsRNA binding motif"
FT Misc-difference 910
FT /note= "Putative zinc coordinating residue"
FT Misc-difference 912
FT /note= "Residue predicted to exhibit proton
FT transfer functions"
FT Misc-difference 966
FT /note= "Putative zinc coordinating residue"
FT Misc-difference 1036
FT /note= "Putative zinc coordinating residue"
FT US5677428-A.
PN XX
XX XX
PD 14-OCT-1997.
XX XX
XX 01-JUN-1995; 95US-0457459.
XX XX
PR 17-FEB-1994; 94US-0197794.
XX 25-JUL-1994; 94US-0280443.
XX XX
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX XX
PI Nishikura K;
XX XX
DR WPI; 1997-511927/47.
DR N-PSDB; AAT89799.
XX XX
PT Antibodies to human double-stranded RNA adenosine deaminase - useful
PT for diagnosis of neurological or CNS disorders, and for drug
PT screening
XX XX
PS Claim 1; Fig 1; 48pp; English.
XX XX
XX This sequence represents a novel double stranded RNA adenosine deaminase
CC (DRADA). DRADA deaminates multiple adenosine residues to inosines by a
CC hydrolytic deamination reaction creating Inosine-Uracil (I-U) mismatched
CC base pairs in double stranded RNA (dsRNA). The accumulation of extensive
CC mismatched I-U base pairs in the dsRNA causes unwinding of the RNA double
CC helix. As the enzyme introduces changes to the sequence of its substrate
CC RNA it is thought to be involved in the RNA editing process. This protein
CC also contains regions which are N-terminal truncated versions of DRADA
CC which are biologically active with similar activity to the full length
CC protein. Three double stranded RNA (dsRNA) binding motifs, DRBM1, DRBM2
CC and DRBM3 are also present and the presence of at least two of these
CC sites are essential for DRADA activity (especially DRBM1 & DRBM3).
CC Antibodies derived from this protein may be used for diagnosis of
CC conditions associated with abnormal, excessive or deficient expression of
CC DRADA e.g. neurological or disorders such as Alzheimer's disease,
CC Huntington's disease, SSPE (subacute sclerosing panencephalitis), measles

CC Inclusion body encephalitis and for drug screening.

XX SQ Sequence 1226 AA;

Query Match 69.6%; Score 35.5; DB 18; Length 1226;
Best Local Similarity 66.7%; Pred. NO. 8.5e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 1; Caps 1;

OY 1 CLHR-GNSC 8

Db 622 CMHKLGNSC 630

Search completed: July 16, 2003, 13:12:40

Job time : 28.3171 secs

PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:10:13 ; Search time 9.36585 Seconds
(without alignments)
25.132 Million cell updates/sec

Title: US-09-910-582B-9
Perfect score: 51
Sequence: 1 CLHRGNSC 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2.6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2.6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2.6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2.6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2.6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2.6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	8	US-09-326-718-9	Sequence 9, Appli
2	36	70.6	99	US-08-985-950-4	Sequence 4, Appli
3	35.5	69.6	72	US-08-280-443-4	Sequence 4, Appli
4	35.5	69.6	72	US-08-457-459-4	Sequence 4, Appli
5	35.5	69.6	72	US-08-555-678-4	Sequence 4, Appli
6	35.5	69.6	72	PCT-US95-02275-4	Sequence 4, Appli
7	35.5	69.6	1226	US-08-280-443-2	Sequence 2, Appli
8	35.5	69.6	1226	US-08-457-459-2	Sequence 2, Appli
9	35.5	69.6	1226	US-08-555-678-2	Sequence 2, Appli
10	35.5	69.6	1226	PCT-US95-02275-2	Sequence 2, Appli
11	35	68.6	3111	US-08-460-309-4	Sequence 4, Appli
12	35	68.6	3111	US-08-125-077-4	Sequence 4, Appli
13	34	66.7	34	US-08-810-009-25	Sequence 25, Appli
14	34	66.7	385	US-09-587-752-4	Sequence 2, Appli
15	34	66.7	386	US-09-058-725B-4	Sequence 4, Appli
16	34	66.7	386	US-09-232-857-4	Sequence 4, Appli
17	34	66.7	389	US-08-789-354-2	Sequence 2, Appli
18	34	66.7	389	US-09-110-937-2	Sequence 2, Appli
19	34	66.7	389	US-09-058-725B-2	Sequence 2, Appli
20	34	66.7	389	US-09-232-857-2	Sequence 2, Appli
21	33	64.7	26	US-09-288-143-159	Sequence 159, App
22	33	64.7	77	US-09-246-500B-2	Sequence 2, Appli
23	33	64.7	345	US-09-040-220D-2	Sequence 2, Appli
24	33	64.7	345	US-09-457-066-2	Sequence 2, Appli
25	33	64.7	345	US-09-457-066-43	Sequence 43, Appli
26	33	64.7	345	US-09-265-686-2	Sequence 2, Appli
27	33	64.7	345	US-09-540-224-5	Sequence 5, Appli

28 62.7 1298 1 US-08-222-616-33 Sequence 33, Appli
29 62.7 1298 1 US-08-340-011-2 Sequence 2, Appli
30 62.7 1298 3 US-08-901-710-2 Sequence 2, Appli
31 62.7 1298 4 US-08-446-648-33 Sequence 33, Appli
32 62.7 1298 5 PCT-US95-04228-33 Sequence 33, Appli
33 62.7 1362 2 US-08-874-678-33 Sequence 33, Appli
34 62.7 1362 3 US-08-643-839-33 Sequence 33, Appli
35 62.7 1362 4 US-09-348-886-33 Sequence 33, Appli
36 62.7 1363 1 US-08-340-011-4 Sequence 4, Appli
37 62.7 1363 2 US-08-874-678-32 Sequence 32, Appli
38 62.7 1363 3 US-08-643-839-32 Sequence 32, Appli
39 62.7 1363 3 US-08-901-710-4 Sequence 4, Appli
40 62.7 1363 4 US-09-348-886-32 Sequence 32, Appli
41 62.7 1368 2 US-08-874-678-34 Sequence 34, Appli
42 62.7 1368 3 US-08-643-839-34 Sequence 34, Appli
43 62.7 1368 4 US-09-348-886-34 Sequence 34, Appli
44 62.7 2471 1 US-08-185-432-16 Sequence 16, Appli
45 62.7 2471 1 US-08-083-590A-19 Sequence 19, Appli

ALIGNMENTS

RESULT 1
US-09-326-718-9
; Sequence 9, Application US/09326718
; Patent No. 6303573
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Mackenna, Deidre A.
; TITLE OF INVENTION: Heart Homing Peptides and Methods of
; TITLE OF INVENTION: Using Same
; FILE REFERENCE: P-LJ 3512
; CURRENT APPLICATION NUMBER: US/09/326,718
; CURRENT FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-09-326-718-9

Query Match 100.0%; Score 51; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
| | | | | | | |

Db 1 CLHRGNSC 8
| | | | | | | |

RESULT 2
US-08-985-950-4
; Sequence 4, Application US/08985950
; Patent No. 6140076
; GENERAL INFORMATION:
; APPLICANT: Adema, Gosse Jan
; TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes;
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

us-09-910-582b-9.rai

Wed Jul 16 13:41:37 2003

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/985,950
;   FILING DATE: 05-DEC-1997
;   CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 60/041,279
;   FILING DATE: 21-MARCH-1997
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 60/033,181
;   FILING DATE: 16-DEC-1996
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 60/032,252
;   FILING DATE: 06-DEC-1996
; ATTORNEY/AGENT INFORMATION:
;   NAME: Ching, Edwin P.
;   REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0670K
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650)852-9196
;   TELEFAX: (650)496-1204
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 99 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-985-950-4
;
; Query Match          70.6%; Score 36; DB 4; Length 99;
; Best Local Similarity 85.7%; Pred. No. 19;
; Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 0;
;
; QY 1 CLHRGNS 7
; DB 12 CLHAGNS 18
;
; RESULT 3
; US-08-280-443-4
; Sequence 4, Application US/08280443
; Patent No. 5643778
; GENERAL INFORMATION:
; APPLICANT: Nishikura, Kazuko
; TITLE OF INVENTION: RNA Editing Enzyme and Methods of Use
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/280,443
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,794
; FILING DATE: 17-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/280,443
; FILING DATE: 25-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST49CUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 72 amino acids
;   TYPE: amino acid
;   TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-457-459-4
;
; Query Match          69.6%; Score 35.5; DB 1; Length 72;
; Best Local Similarity 66.7%; Pred. No. 17;
; Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 1;
;
; QY 1 CLHR-GNSC 8
; DB 10 CMHKLGNCS 18
;
; RESULT 4
; US-08-457-459-4
; Sequence 4, Application US/08457459
; Patent No. 5677428
; GENERAL INFORMATION:
; APPLICANT: Nishikura, Kazuko
; TITLE OF INVENTION: RNA Editing Enzyme and Methods of Use
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P.O. Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,459
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,794
; FILING DATE: 17-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/280,443
; FILING DATE: 25-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST49CUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 72 amino acids
;   TYPE: amino acid
;   TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-457-459-4
;
; Query Match          69.6%; Score 35.5; DB 1; Length 72;
; Best Local Similarity 66.7%; Pred. No. 17;
; Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 1;
;
; QY 1 CLHR-GNSC 8
; DB 10 CMHKLGNCS 18
;
; RESULT 5
```

US-08-555-678-4

; Sequence 4, Application US/08555678
; Patent No. 5763174
; GENERAL INFORMATION:
; APPLICANT: Nishikura, Kazuko
; TITLE OF INVENTION: RNA Editing Enzyme and Methods
; TITLE OF INVENTION: Of Use Thereof
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P.O. Box 457
; City: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/555,678
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,794
; FILING DATE: 17-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/280,443
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,459
; FILING DATE: 01-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST49DUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 72 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US95-678-4

Query Match 69.6%; Score 35.5; DB 1; Length 72;

Best Local Similarity 66.7%; Pred. No. 17;

Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Qy 1 CLHR-GNSC 8

I:::||||

Db 10 CMHKLNSC 18

RESULT 6

PCT-US95-02275-4

; Sequence 4, Application PC/TUS9502275
; GENERAL INFORMATION:
; APPLICANT: Wistar Institute of Anatomy & Biology
; TITLE OF INVENTION: RNA Editing Enzyme and Methods of Use
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P.O. Box 457
; City: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02275
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/280,443
; FILING DATE: 25-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,794
; FILING DATE: 17-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST49BPCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 72 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US95-02275-4

Query Match 69.6%; Score 35.5; DB 5; Length 72;

Best Local Similarity 66.7%; Pred. No. 17;

Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Qy 1 CLHR-GNSC 8

I:::||||

Db 10 CMHKLNSC 18

RESULT 7

US-08-280-443-2

; Sequence 2, Application US/08280443
; Patent No. 5643778
; GENERAL INFORMATION:
; APPLICANT: Nishikura, Kazuko
; TITLE OF INVENTION: RNA Editing Enzyme and Methods of Use
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P.O. Box 457
; City: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/280,443
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,794
; FILING DATE: 17-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST49AUSA
; TELECOMMUNICATION INFORMATION:

us-09-910-582b-9.ra1

Wed Jul 16 13:41:37 2003

TELEPHONE: 215-540-9206
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1226 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-280-443-2

Query Match 69.6%; Score 35.5; DB 1; Length 1226;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 CLHR-GNSC 8
Db 622 CMHKLGNCS 630

RESULT 8
US-08-457-459-2

Sequence 2, Application US/08457459
Patent No. 5677428

GENERAL INFORMATION:

APPLICANT: Nishikura, Kazuko
TITLE OF INVENTION: RNA Editing Enzyme and Methods of Use
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P.O. Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/457,459

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/197,794

FILING DATE: 17-FEB-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/280,443

FILING DATE: 25-JUL-1994

ATTORNEY/AGENT INFORMATION:

NAME: Bak, Mary E.

REGISTRATION NUMBER: 31,215

REFERENCE/DOCKET NUMBER: WST49CUSA

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-540-9206

TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1226 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-457-459-2

Query Match 69.6%; Score 35.5; DB 1; Length 1226;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 CLHR-GNSC 8
Db 622 CMHKLGNCS 630

RESULT 9

US-08-555-678-2

Sequence 2, Application US/08555678

Patent No. 5763174

GENERAL INFORMATION:

APPLICANT: Nishikura, Kazuko

TITLE OF INVENTION: RNA Editing Enzyme and Methods

TITLE OF INVENTION: of Use Thereof

NUMBER OF SEQUENCES: 67

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Cntr, P.O. Box 457

CITY: Spring House

STATE: Pennsylvania

COUNTRY: USA

ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/555,678

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/197,794

FILING DATE: 17-FEB-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/280,443

FILING DATE: 25-JUL-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/457,459

FILING DATE: 01-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Bak, Mary E.

REGISTRATION NUMBER: 31,215

REFERENCE/DOCKET NUMBER: WST49DUSA

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-540-9206

TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1226 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-555-678-2

Query Match 69.6%; Score 35.5; DB 1; Length 1226;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 CLHR-GNSC 8
Db 622 CMHKLGNCS 630

RESULT 10

PCT-US95-02275-2

Sequence 2, Application PC/TUS9502275

GENERAL INFORMATION:

APPLICANT: Wistar Institute of Anatomy & Biology

TITLE OF INVENTION: RNA Editing Enzyme and Methods of Use

TITLE OF INVENTION: Thereof

NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Cntr, P.O. Box 457

CITY: Spring House

STATE: Pennsylvania

COUNTRY: USA

ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02275
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/280,443
FILING DATE: 25-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/197,794
FILING DATE: 17-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: WST49BPCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9206
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1226 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-02275-2

Query Match 69.6%; Score 35.5; DB 5; Length 1226;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 CLHR-GNSC 8
|:|: ||||
Db 622 CMHKLGNCS 630

RESULT 11
US-08-460-309-4
; Sequence 4, Application US/08460309
; Patent No. 5837496
; GENERAL INFORMATION:
; APPLICANT: Engvall, Eva
; APPLICANT: Leivo, Ilmo
; TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
; OPERATING SYSTEM: PC-DOS/MS-DOS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,309
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/125,077
; FILING DATE: 22-SEP-1993
; APPLICATION NUMBER: US PCT/US 94/10730
; FILING DATE: 21-SEP-1994
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3111 amino acids

FILING DATE: 30-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/919,951
FILING DATE: 27-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 9721
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3111 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-460-309-4
Query Match 68.6%; Score 35; DB 2; Length 3111;
Best Local Similarity 62.5%; Pred. No. 6.9e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 CLHRGNSC 8
|:|: ||||
Db 1016 CSHLGNCS 1023
RESULT 12
US-08-125-077-4
; Sequence 4, Application US/08125077
; Patent No. 5872231
; Patent No. 5872231 5840863
; GENERAL INFORMATION:
; APPLICANT: Engvall, Eva
; APPLICANT: Leivo, Ilmo
; TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
; OPERATING SYSTEM: PC-DOS/MS-DOS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/125,077
; FILING DATE: 22-SEP-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US 94/10730
; FILING DATE: 21-SEP-1994
; APPLICATION NUMBER: US 07/472,319
; FILING DATE: 30-JAN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/919,951
; FILING DATE: 27-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 9721
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3111 amino acids

us-09-910-582b-9.ra1

Wed Jul 16 13:41:37 2003

```

;
; TYPE: amino acid
; TOPOLOGY: linear
US-08-125-077-4
Query Match 68.6%; Score 35; DB 2; Length 3111;
Best Local Similarity 62.5%; Pred. No. 6.9e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
Db 1016 CSHLGNNC 1023

RESULT 13
US-08-810-009-25
; Sequence 25, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; APPLICANT: Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P. O. Drawer 34009
; CITY: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/810,009
; APPLICATION NUMBER: 04-MAR-1997
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-810-009-25
Query Match 66.7%; Score 34; DB 4; Length 34;
Best Local Similarity 71.4%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CLHRGNS 7
Db 3 CMHRGTS 9

RESULT 14
US-09-587-754-2
; Sequence 2, Application US/09587754
; Patent No. 6428982
; GENERAL INFORMATION:
; APPLICANT: Nabil A. Elshourbagy
; APPLICANT: Usman Shabon
; APPLICANT: Derk J. Bergsma
; TITLE OF INVENTION: MOUSE UROTENSIN-II RECEPTOR (UTB-R)
; FILE REFERENCE: GH-70626
; CURRENT APPLICATION NUMBER: US/09/587,754
; CURRENT FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: 60/137,834
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 385
; TYPE: PRT
; ORGANISM: MUS MUSCULUS
US-09-587-754-2
Query Match 66.7%; Score 34; DB 4; Length 385;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
Db 303 CLTYGNSC 310

RESULT 15
US-09-058-725B-4
; Sequence 4, Application US/09058725B
; Patent No. 6133420
; GENERAL INFORMATION:
; APPLICANT: Ames, Robert
; APPLICANT: Sarau, Henry
; APPLICANT: Foley, James
; APPLICANT: Chamber, Jon
; TITLE OF INVENTION: A Method of Finding Angonist
; TITLE OF INVENTION: and Antagonist To Human and Rat GPR 14
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/058,725B
; FILING DATE: April 10, 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/789,354
; FILING DATE: 27-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T.
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: GP50005-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 386 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-058-725B-4
Query Match 66.7%; Score 34; DB 4; Length 386;
; APPLICANT: Nabil A. Elshourbagy
; APPLICANT: Usman Shabon
; APPLICANT: Derk J. Bergsma
; TITLE OF INVENTION: MOUSE UROTENSIN-II RECEPTOR (UTB-R)
; FILE REFERENCE: GH-70626
; CURRENT APPLICATION NUMBER: US/09/587,754
; CURRENT FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: 60/137,834
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 385
; TYPE: PRT
; ORGANISM: MUS MUSCULUS
US-09-587-754-2
```


Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CLHRGNSC 8
Db 304 CLTYGNSC 311

Search completed: July 16, 2003, 13:17:08
Job time : 10.3659 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:15:24 ; Search time 15.2195 Seconds
(without alignments)
62.425 Million cell updates/sec

Title: US-09-910-582B-9
Perfect score: 51
Sequence: 1 CLHRGNSC 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	8	9	US-09-910-582B-9
2	51	100.0	8	10	US-09-782-650-4
3	36	70.6	99	9	US-10-290-631-4
4	36	70.6	224	9	US-09-866-050A-711
5	36	70.6	4636	10	US-09-835-996A-33
6	35.5	69.6	1225	9	US-10-233-553-20
7	35	68.6	555	9	US-09-927-827-49
8	35	68.6	920	10	US-09-934-868-52
9	35	68.6	1354	10	US-09-808-571A-4
10	35	68.6	1447	10	US-09-808-571A-2
11	35	68.6	1905	9	US-10-270-333-9
12	35	68.6	3070	9	US-09-961-403-7
13	34	66.7	34	10	US-09-776-490-25
14	34	66.7	34	10	US-09-776-491-25
15	34	66.7	389	9	US-10-225-567A-538
16	33	64.7	26	9	US-10-150-111-159
17	33	64.7	28	10	US-09-864-761-39799
18	33	64.7	37	10	US-09-894-882-223
19	33	64.7	37	10	US-09-894-882-387

20	64.7	46	10	US-09-864-761-44167	Sequence 44167, A
21	64.7	73	10	US-09-894-882-222	Sequence 222, App
22	64.7	88	9	US-10-086-623-11	Sequence 11, Appl
23	64.7	88	9	US-10-260-539-11	Sequence 11, Appl
24	64.7	182	9	US-09-852-209A-16	Sequence 16, Appl
25	64.7	182	9	US-09-852-209A-17	Sequence 17, Appl
26	64.7	182	9	US-10-131-600-16	Sequence 16, Appl
27	64.7	182	9	US-10-131-600-17	Sequence 17, Appl
28	64.7	318	9	US-09-852-209A-5	Sequence 5, Appl1
29	64.7	318	9	US-10-131-600-5	Sequence 5, Appl1
30	64.7	339	10	US-09-925-302-776	Sequence 776, App
31	64.7	345	9	US-09-978-295A-488	Sequence 488, App
32	64.7	345	9	US-09-852-209A-3	Sequence 3, Appl1
33	64.7	345	9	US-09-852-209A-7	Sequence 7, Appl1
34	64.7	345	9	US-10-086-623-32	Sequence 32, Appl
35	64.7	345	9	US-09-978-697-488	Sequence 488, App
36	64.7	345	9	US-09-978-192A-488	Sequence 488, App
37	64.7	345	9	US-10-139-583-2	Sequence 2, Appl1
38	64.7	345	9	US-10-139-583-43	Sequence 43, Appl
39	64.7	345	9	US-09-999-832A-488	Sequence 488, App
40	64.7	345	9	US-09-978-189-488	Sequence 488, App
41	64.7	345	9	US-10-028-072-286	Sequence 286, App
42	64.7	345	9	US-10-121-049-286	Sequence 286, App
43	64.7	345	9	US-10-123-904-286	Sequence 286, App
44	64.7	345	9	US-10-140-470-286	Sequence 286, App
45	64.7	345	9	US-09-796-753-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-09-910-582B-9
; Sequence 9, Application US/09910582B
; Publication No. US20030045476A1
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; TITLE OF INVENTION: Heart Homing Conjugates
; FILE REFERENCE: P-LJ 4857
; CURRENT APPLICATION NUMBER: US/09/910,582B
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 09/326,718
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; TYPE: PRT
; LENGTH: 8
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-09-910-582B-9

Query Match 100.0%; Score 51; DB 9; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CLHRGNSC 8
| | | | | | | |
Db 1 CLHRGNSC 8

RESULT 2
US-09-782-650-4
; Sequence 4, Application US/09782650
; Patent No. US20020019350A1
; GENERAL INFORMATION:
; APPLICANT: Levine, Arnold J.
; APPLICANT: Mitterer, Artur
; APPLICANT: Falkner, Falko-Guenter
; APPLICANT: Scheiflinger, Friedrich
; APPLICANT: Dörner, Friedrich

Wed Jul 16 13:41:37 2003

APPLICANT: Edwards Lifesciences Corporation
TITLE OF INVENTION: Targeted Angiogenesis
FILE REFERENCE: 2053D-000611US
CURRENT APPLICATION NUMBER: US/09/782.650
CURRENT FILING DATE: 2001-02-12
PRIOR APPLICATION NUMBER: US 09/324,079
PRIOR FILING DATE: 1999-06-01
PRIOR APPLICATION NUMBER: US 09/327,045
PRIOR FILING DATE: 1999-06-07
PRIOR APPLICATION NUMBER: PCT/US00/14988
PRIOR FILING DATE: 2000-05-31
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO. 4
LENGTH: 8
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:targeting
OTHER INFORMATION: molecule
US-09-782-650-4
Query Match 100.0%; Score 51; DB 10; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CLHRGNSC 8
DB 1 CLHRGNSC 8
RESULT 3
US-10-290-631-4
Sequence 4, Application US/10290631
Publication No. US20030105303A1
GENERAL INFORMATION:
APPLICANT: Adema, Gosse Jan
Meynard, Linde
Gorman, Daniel M.
McClanahan, Terrill K.
Zurawski, Sandra M.
Zurawski, Gerard
Lanier, Lewis L.
Phillips Jr., Joseph H.
TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes;
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: DNAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/290,631
FILING DATE: 08-NOV-2003
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/985,950
FILING DATE: 05-DEC-1997
APPLICATION NUMBER: US 60/041,279
FILING DATE: 21-MARCH-1997
APPLICATION NUMBER: US 60/033,181
FILING DATE: 16-DEC-1996
APPLICATION NUMBER: US 60/032,252
FILING DATE: 06-DEC-1996
ATTORNEY/AGENT INFORMATION:

NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: DX0670K
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)852-9196
TELEFAX: (650)496-1204
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 99 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-290-631-4
Query Match 70.6%; Score 36; DB 9; Length 99;
Best Local Similarity 85.7%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CLHRGNS 7
DB 12 CLHAGNS 18
RESULT 4
US-09-866-050A-711
Sequence 711, Application US/09866050A
Publication No. US2003004071A1
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James G.
TITLE OF INVENTION: Compositions Isolated From Skin Cells
TITLE OF INVENTION: and Methods for Their Use
FILE REFERENCE: 11000.1011c4U
CURRENT APPLICATION NUMBER: US/09/866.050A
CURRENT FILING DATE: 2001-05-24
NUMBER OF SEQ ID NOS: 725
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 711
LENGTH: 224
TYPE: PRT
ORGANISM: Mouse
US-09-866-050A-711
Query Match 70.6%; Score 36; DB 9; Length 224;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CLHRGNS 7
DB 25 CLHAGNS 31
RESULT 5
US-09-835-996A-33
Sequence 33, Application US/09835996A
Patent No. US20020142953A1
GENERAL INFORMATION:
APPLICANT: Ballinger, Dennis
APPLICANT: Loeb, Debra
APPLICANT: Montgomery, Julie
APPLICANT: Tang, Y. Tom
APPLICANT: Goodrich, Ryle
APPLICANT: Liu, Chinghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhao, Qing
APPLICANT: Wehrman, Tom
APPLICANT: Drmanac, Radoje



; APPLICANT: Ren, Feiyan
; APPLICANT: Qian, Xiaohong
; APPLICANT: Wang, Dunrui
; TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
; FILE REFERENCE: 28110/35915A
; CURRENT APPLICATION NUMBER: US/09/835,996A
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: US 60/197,137
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US 09/714,936
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: US 09/667,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33
; LENGTH: 4636
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-835-996A-33

Query Match 70.6%; Score 36; DB 10; Length 4636;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CLHRGNS 7
||:||||
DB 2195 CLYGRNS 2201

RESULT 6

US-10-233-553-20
; Sequence 20, Application US/10233553
; Publication No. US20030125285A1
; GENERAL INFORMATION:
; APPLICANT: NIPPON SHINYAKU CO., LTD.
; APPLICANT: HIRABAYASHI, Kazuko
; APPLICANT: YANO, Junichi
; TITLE OF INVENTION: Method of examining the efficacy of therapy with nucleic acid
; FILE REFERENCE: B-345
; CURRENT APPLICATION NUMBER: US/10/233,553
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: JP 2001-267385
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 31
; SEQ ID NO 20
; LENGTH: 1225
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-233-553-20

Query Match 69.6%; Score 35.5; DB 9; Length 1225;
Best Local Similarity 66.7%; Pred. No. 5.7e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 CLHR-GNSC 8
|:|:||||
DB 621 CMHKLGNSC 629

RESULT 7

US-09-927-827-49
; Sequence 49, Application US/09927827
; Publication No. US20030036176A1
; GENERAL INFORMATION:
; APPLICANT: Bower, Stanley G.
; APPLICANT: Ramseier, Thomas M.
; TITLE OF INVENTION: Directed Genetic Engineering of Xanthomonas campestris
; FILE REFERENCE: 38-10(15824)B

; CURRENT APPLICATION NUMBER: US/09/927,827
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/279,493
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 69
; SEQ ID NO 49
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-09-927-827-49

Query Match 68.6%; Score 35; DB 9; Length 555;
Best Local Similarity 85.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LHRGNSC 8
|||||
DB 350 LHRGNMC 356

RESULT 8

US-09-934-868-52
; Sequence 52, Application US/09934868
; Patent No. US20020137190A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, James M
; APPLICANT: Schenzle, Andreas J
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1596 US NA
; CURRENT APPLICATION NUMBER: US/09/934,868
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 52
; LENGTH: 920
; TYPE: PRT
; ORGANISM: Methylomonas 16a
; FEATURE:
; OTHER INFORMATION: Nasa
US-09-934-868-52

Query Match 68.6%; Score 35; DB 10; Length 920;
Best Local Similarity 62.5%; Pred. No. 5.4e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
|||||
DB 869 CLHAGTGC 876

RESULT 9

US-09-808-571A-4
; Sequence 4, Application US/09808571A
; Patent No. US20020106723A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; TITLE OF INVENTION: Receptor for latrotoxin from insects
; FILE REFERENCE: Le A 34 402
; CURRENT APPLICATION NUMBER: US/09/808,571A
; CURRENT FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: DE 100 13 580.3
; PRIOR FILING DATE: 2000-03-18
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1354
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-808-571A-4

us-09-910-582b-9.rapb

Wed Jul 16 13:41:37 2003

```
Query Match      68.6%; Score 35; DB 10; Length 1354;
Best Local Similarity 62.5%; Pred. No. 7.5e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 CLHRGNSC 8
DB      402 CLHNSSC 409

RESULT 10
US-09-808-571A-2
; Sequence 2, Application US/09808571A
; Patent No. US20020106723A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; TITLE OF INVENTION: Receptor for latrotoxin from insects
; FILE REFERENCE: Le A 34 402
; CURRENT APPLICATION NUMBER: US/09/808,571A
; CURRENT FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: DE 100 13 580.3
; PRIOR FILING DATE: 2000-03-18
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1447
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-808-571A-2

Query Match      68.6%; Score 35; DB 10; Length 1447;
Best Local Similarity 62.5%; Pred. No. 8e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 CLHRGNSC 8
DB      403 CLHNSSC 410

RESULT 11
US-10-270-333-9
; Sequence 9, Application US/10270333
; Publication No. US20030092124A1
; GENERAL INFORMATION:
; APPLICANT: Craychik, Anibal
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS, AND USES
; FILE REFERENCE: THEREOF AS INSECTICIDAL TARGETS
; CURRENT APPLICATION NUMBER: US/10/270,333
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/168,677
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/175,691
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/191,638
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1905
; TYPE: PRT
; ORGANISM: Drosophila
US-10-270-333-9

Query Match      68.6%; Score 35; DB 9; Length 1905;
Best Local Similarity 62.5%; Pred. No. 1e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 CLHRGNSC 8
DB      395 CLHNSSC 402
```

```
RESULT 12
US-09-961-403-7
; Sequence 7, Application US/09961403
; Publication No. US20030077589A1
; GENERAL INFORMATION:
; APPLICANT: HE-STUMPP, HOLGER
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: KRAETZSCHMAR, JOERN
; APPLICANT: KREFT, BERTHOLT
; APPLICANT: WINTERHAGER, ELKE
; APPLICANT: REGIDOR, PEDRO
; APPLICANT: SCOTTI, SIMONE
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
; FILE REFERENCE: SCH-1789
; CURRENT APPLICATION NUMBER: US/09/961,403
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 3070
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-961-403-7

Query Match      68.6%; Score 35; DB 9; Length 3070;
Best Local Similarity 62.5%; Pred. No. 1.5e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 CLHRGNSC 8
DB      1016 CSHLGNNC 1023

RESULT 13
US-09-776-490-25
; Sequence 25, Application US/09776490
; Patent No. US20010012886A1
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; APPLICANT: Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: No. US20010012886A1th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/776,490
; FILING DATE: 02-Feb-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/810,009
; FILING DATE: 04-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
```

LENGTH: 34 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-776-490-25

Query Match 66.7%; Score 34; DB 10; Length 34;
Best Local Similarity 71.4%; Pred. No. 43;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CLHRGNS 7
I:||||
Db 3 CMHRGTS 9

RESULT 14

US-09-776-491-25
Sequence 25, Application US/09776491
Patent No. US20010013135A1

GENERAL INFORMATION:

APPLICANT: Briggs, Steven P.
Johal, John
Gray, John

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
CELL DEATH AND DISEASE RESISTANCE IN PLANTS

NUMBER OF SEQUENCES: 65

CORRESPONDENCE ADDRESSES:

ADDRESSEE: BELL, SELTZER, PARK & GIBSON

STREET: P.O. Drawer 34009

CITY: Charlotte

STATE: No. US20010013135A1th Carolina

COUNTRY: USA

ZIP: 28234

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/776.491

FILING DATE: 02-Feb-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/810,009

FILING DATE: 04-MAR-1997

ATTORNEY/AGENT INFORMATION:

NAME: Spruill, W. Murray

REGISTRATION NUMBER: 32,943

REFERENCE/DOCKET NUMBER: 5718-4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-881-3140

TELEFAX: 919-881-3175

TELEX: 575102

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 34 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 25:

US-09-776-491-25

Query Match 66.7%; Score 34; DB 10; Length 34;
Best Local Similarity 71.4%; Pred. No. 43;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CLHRGNS 7
I:||||
Db 3 CMHRGTS 9

RESULT 15

US-10-225-567A-538

Sequence 538, Application US/10225567A

Publication No. US20030113798A1

GENERAL INFORMATION:

APPLICANT: Lifespan Biosciences

APPLICANT: Brown, Joseph P.

APPLICANT: Burmer, Glenna C.

APPLICANT: Roush, Christine L.

TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPT

FILE REFERENCE: 1920-4-4

CURRENT APPLICATION NUMBER: US/10/225.567A

CURRENT FILING DATE: 2001-12-19

PRIOR APPLICATION NUMBER: 60/257,144

PRIOR FILING DATE: 2000-12-19

NUMBER OF SEQ ID NOS: 2292

SOFTWARE: Patentin version 3.1

SEQ ID NO 538

LENGTH: 389

TYPE: PRT

ORGANISM: Homo sapiens

US-10-225-567A-538

Query Match 66.7%; Score 34; DB 9; Length 389;

Best Local Similarity 75.0%; Pred. No. 3.7e+02;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CLHRGNSC 8

I:||||

Db 302 CLTYGNSC 309

Search completed: July 16, 2003, 13:35:59

Job time : 16.2195 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:09:18 ; Search time 10.5366 seconds
(without alignments)
72.991 Million cell updates/sec

Title: US-09-910-582B-9

Perfect score: 51

Sequence: 1 CLHRGNSC 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: p1r1.*

2: p1r2.*

3: p1r3.*

4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	76.5	631	2	D96683
2	38	74.5	1743	2	T26859
3	37	72.5	452	2	C37162
4	36.5	71.6	524	2	T40844
5	36	70.6	1225	2	A56514
6	35.5	69.6	1175	2	I57549
7	35.5	69.6	1226	1	S65593
8	35	68.6	177	2	T43665
9	35	68.6	192	2	T15719
10	35	68.6	207	2	A56385
11	35	68.6	418	2	T16713
12	35	68.6	423	2	T21570
13	35	68.6	440	2	T21568
14	35	68.6	1133	1	EGRT
15	35	68.6	3106	1	S38868
16	34	66.7	60	2	S38335
17	34	66.7	82	2	S58530
18	34	66.7	196	2	S49194
19	34	66.7	299	2	C89254
20	34	66.7	386	2	I84612
21	34	66.7	409	2	S44171
22	34	66.7	424	2	S43560
23	34	66.7	745	2	T03119
24	33	64.7	77	2	A01259
25	33	64.7	142	2	A82517
26	33	64.7	147	2	E59780
27	33	64.7	221	2	E38868
28	33	64.7	245	2	T49889
29	33	64.7	301	2	T32803

30 33 64.7 414 2 S65200
31 33 64.7 450 2 T31256
32 33 64.7 459 2 T31134
33 33 64.7 489 2 C70707
34 33 64.7 552 2 T30972
35 33 64.7 553 2 B98207
36 33 64.7 553 2 AG3079
37 33 64.7 1161 2 T18642
38 32 62.7 60 2 JC2419
39 32 62.7 79 2 B84252
40 32 62.7 89 2 E97731
41 32 62.7 127 2 D64545
42 32 62.7 127 2 E71962
43 32 62.7 157 2 T44794
44 32 62.7 191 2 T19573
45 32 62.7 202 2 T33336

ALIGNMENTS

RESULT 1

D96683

hypothetical protein FL2P19.8 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: D96683

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Hulzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: AB6141; MUID:21016719; PMID:11130712

A:Accession: D96683

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-631 <STO>

A:Cross-references: GB:AE005173; NID:g6227016; PIDN:AAF06052.1; GSPDB:GN00141

C:Genetics:

A:Gene: FL2P19.8

A:Map position: 1

Query Match

Best Local Similarity 76.5% Score 39; DB 2; Length 631;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CLHRGNS 7

DB 547 CVHRGNS 553

RESULT 2

T26859

hypothetical protein Y43F8B.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T26859

R:Ainscough, R.

submitted to the EMBL Data Library, October 1998

A:Reference number: Z20278

A:Accession: T26859

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-1743 <WIL>

A:Cross-references: EMBL:AL032623; PIDN:CAA21511.1; CESP:Y43F8B.3

A:Experimental source: clone Y43F8B

C:Genetics:

A:Gene: CESP:Y43F8B.3

Wed Jul 16 13:41:38 2003

us-09-910-582b-9.rpr

```
C:Species: Gallus gallus (chicken)
C:Date: 21-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 02-Feb-2001
C:Accession: A56514, 150691
R:Wang, S.Z.; Adler, R.
J. Cell Biol. 128, 761-768, 1995
A:Title: Chromokinesin: a DNA-binding, kinesin-like nuclear protein.
A:Reference number: A56514; MUID:95181533; PMID:7876303
A:Accession: A56514
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1225 <WAK>
A:Cross-references: GB:U18309; NID:9603760; PIDN:AAC59666.1; PID:g603761
R:Wang, S.Z.; Adler, R.
Proc. Natl. Acad. Sci. U.S.A. 91, 1351-1355, 1994
A:Title: A developmentally regulated basic-leucine zipper-like gene and its expression
A:Reference number: A53451; MUID:94151328; PMID:8108415
A:Accession: 150691
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 728-1086, RI' <WA2>
A:Cross-references: EMBL:U04821; NID:g440792; PIDN:AAA18960.1; PID:g440793
C:Genetics:
A:Gene: sw3-3
C:Superfamily: unassigned kinesin-related proteins; kinesin motor domain homology
C:Keywords: ATP; DNA binding; mitosis; nucleotide binding; P-loop
F:11-344/Domain: kinesin motor domain homology <KWOT>
F:88-95/Region: nucleotide-binding motif A (P-loop)

Query Match 70.6%; Score 36; DB 2; Length 1225;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
DB 191 CLEOGNNC 198

RESULT 6
I57549
adenosine deaminase (EC 3.5.4.4) double-stranded RNA-specific - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 22-Jun-1999
C:Accession: I57549
R:O'Connell, M.A.; Krause, S.; Higuchi, M.; Hsuan, J.J.; Totty, N.F.; Jenny, A.; Kell
Mol. Cell. Biol. 15, 1389-1397, 1995
A:Title: Cloning of cDNAs encoding mammalian double-stranded RNA-specific adenosine d
A:Reference number: I57549; MUID:95166224; PMID:7862132
A:Accession: I57549
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1175 <RES>
A:Cross-references: EMBL:U18942; NID:g755816; PIDN:AAA5039.1; PID:g755817
C:Superfamily: double-stranded RNA-specific adenosine deaminase; double-stranded RNA-
C:Keywords: hydrolase; metalloprotein; RNA binding; RNA editing; zinc
F:452-522/Domain: double-stranded RNA-binding repeat homology <DSR1>
F:561-633/Domain: double-stranded RNA-binding repeat homology <DSR2>
F:670-741/Domain: double-stranded RNA-binding repeat homology <DSR3>
F:856,912,982/Binding site: zinc (His, Cys, Cys) #status predicted

Query Match 69.6%; Score 35.5; DB 2; Length 1175;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 CLHRGNSC 8
DB 572 CMHKLGNCS 580

RESULT 7
S65593
adenosine deaminase (EC 3.5.4.4), double-stranded RNA-specific - human
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
```

```
A: Introns: 65/1; 92/2; 128/1; 229/1; 367/1; 422/1; 486/2; 523/1; 571/1; 628/1; 857/2; 94
Query Match 74.5%; Score 38; DB 2; Length 1743;
Best Local Similarity 75.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
DB 1557 CARRGNSC 1564

RESULT 3
C97162
UDP-N-acetylmuramyl pentapeptide synthase [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: C97162
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A:Reference number: A96500; MUID:21359325; PMID:21359325
A:Accession: C97162
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-452 <KUR>
A:Cross-references: GB:AB001437; PIDN:AAK80086.1; PID:g15025118; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC2128

Query Match 72.5%; Score 37; DB 2; Length 452;
Best Local Similarity 85.7%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CLHRGNS 7
DB 256 CLHEGNS 262

RESULT 4
T40844
probable mitochondrial translation system component - fission yeast (Schizosaccharomyces
C:Species: Schizosaccharomyces pombe
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C:Accession: T40844
R:Hilbert, H.; Duesterhoeft, A.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, September 1998
A:Reference number: Z21951
A:Accession: T40844
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-524 <HIL>
A:Cross-references: EMBL:AL031740; PIDN:CAA21084.1; GSPDB:GN00068; SPDB:SPCC1183.04c
A:Experimental source: strain 972h; cosmid c1183
C:Genetics:
A:Gene: SPDB:SPCC1183.04c
A:Map position: 3
A:Introns: 48/1
C:Keywords: mitochondrion

Query Match 71.6%; Score 36.5; DB 2; Length 524;
Best Local Similarity 63.6%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 0; Indels 3; Gaps 1;

QY 1 CLHRG---NSC 8
DB 20 CLHRGFQINSC 30

RESULT 5
A56514
chromokinesin - chicken
```


Query Match 68.6%; Score 35; DB 2; Length 418;
 Best Local Similarity 62.5%; Pred. No. 67;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
 DB 286 CICKGNTC 293
 I: |||I|

RESULT 12
 T21570
 hypothetical protein F30A10.8b - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 A:Accession: T21570
 R:Barlow, K.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z19442
 A:Accession: T21570
 A:Status: Preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-423 <WIL>
 A:Cross-references: EMBL:Z81072; PIDN:CAB03028.1; GSPDB:GN00019; CESP:F30A10.8b
 A:Experimental source: clone F30A10
 C:Genetics:
 A:Gene: CESP:F30A10.8b
 A:Map position: 1
 A:Introns: 32/3; 73/2; 378/3

Query Match 68.6%; Score 35; DB 2; Length 423;
 Best Local Similarity 75.0%; Pred. No. 68;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
 DB 353 CLWRGESC 360
 I: |||I|I|

RESULT 13
 T21568
 hypothetical protein F30A10.8a - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 A:Accession: T21568
 R:Barlow, K.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z19442
 A:Accession: T21568
 A:Status: Preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-440 <WIL>
 A:Cross-references: EMBL:Z81072; PIDN:CAB03025.1; GSPDB:GN00019; CESP:F30A10.8a
 A:Experimental source: clone F30A10
 C:Genetics:
 A:Gene: CESP:F30A10.8a
 A:Map position: 1
 A:Introns: 32/3; 73/2; 378/3; 410/3

Query Match 68.6%; Score 35; DB 2; Length 440;
 Best Local Similarity 75.0%; Pred. No. 70;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
 DB 353 CLWRGESC 360
 I: |||I|I|

RESULT 14
 EGRT
 epidermal growth factor precursor - rat
 N:Alternate names: urogastone precursor
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-Dec-1988 #sequence_revision 14-Aug-1998 #text_change 18-Jun-1999

C:Accession: I52995; S05074; S01974; A25425; S18419; S08288
 R:Saggi, S.J.; Safirstein, R.; Price, P.M.
 DNA Cell Biol. 11, 481-487, 1992
 A:Title: Cloning and Sequencing of the Rat Preproepidermal Growth Factor cDNA: Compar
 A:Reference number: I52995; MUID:92398779; PMID:1524680
 A:Accession: I52995
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: mRNA
 A:Residues: 1-1133 <RES>
 A:Cross-references: EMBL:U04842; NID:g440236; PIDN:AAB60436.1; PID:g440237
 R:Simpson, R.
 submitted to the EMBL Data Library, August 1988
 A:Reference number: S05074
 A:Accession: S05074
 A:Molecule type: mRNA
 A:Residues: 'W', 966, 'RWL', 970-1023, 'NW', 1026-1108, 'SGAGVSGPQFWFVLE', 1126, 'HQ' <SIM>
 A:Cross-references: EMBL:X12748
 R:Dorow, D.S.; Simpson, R.J.
 Nucleic Acids Res. 16, 9338, 1988
 A:Title: Cloning and sequence analysis of a cDNA for rat epidermal growth factor.
 A:Reference number: S01974; MUID:89016634; PMID:3262867
 A:Accession: S01974
 A:Molecule type: mRNA
 A:Residues: 'W', 966, 'RWL', 970-1023, 'NW', 1026-1108 <DOR>
 A:Cross-references: EMBL:X12748
 R:Simpson, R.J.; Smith, J.A.; Moritz, R.L.; O'Hare, M.J.; Rudland, P.S.; Morrison, J.
 Eur. J. Biochem. 153, 629-637, 1985
 A:Title: Rat epidermal growth factor: complete amino acid sequence.
 A:Reference number: A25425; MUID:86081810; PMID:3000782
 A:Accession: A25425
 A:Molecule type: protein
 A:Residues: 974-1021 <SI>
 R:Nishi, N.; Shimizu, C.; Okutani, T.; Kagawa, Y.; Takasuga, H.; Suno, M.; Wada, F.
 Biochim. Biophys. Acta 1095, 268-275, 1991
 A:Title: Rat prostatic growth factors: purification and characterization of high and
 A:Reference number: S18419; MUID:92069070; PMID:1958699
 A:Accession: S18419
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 974-1021 <NIS>
 R:Nexo, E.; Jorgensen, P.E.; Thim, L.; Roepstorff, P.
 Biochim. Biophys. Acta 1037, 388-393, 1990
 A:Title: Purification and characterization of a low and a high molecular weight form
 A:Reference number: S08288; MUID:90181442; PMID:2310752
 A:Accession: S08288
 A:Molecule type: protein
 A:Residues: 974-1024 <NEX>
 C:Comment: Epidermal growth factor (EGF) stimulates the proliferation and differenti-
 gastrointestinal cell proliferation.
 C:Comment: EGF is released in the pancreas, small intestine, mammary gland, and (in
 C:Comment: The EGF precursor is found in kidney as a receptor-like membrane-bound pro
 C:Superfamily: epidermal growth factor precursor; EGF homology; LDL receptor YWTD-con
 C:Keywords: duplication; growth factor; mitogen; tandem repeat; transmembrane protein
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-1133/Product: epidermal growth factor propiortein, membrane-bound form #status pr
 F:22-1035/Domain: extracellular #status predicted <EXT>
 F:44-480/Region: EGF precursor long repeat <LR1>
 F:47-86/Domain: LDL receptor YWTD-containing repeat homology <YW01>
 F:87-128/Domain: LDL receptor YWTD-containing repeat homology <YW02>
 F:129-170/Domain: LDL receptor YWTD-containing repeat homology <YW03>
 F:171-212/Domain: LDL receptor YWTD-containing repeat homology <YW04>
 F:213-237/Domain: LDL receptor YWTD-containing repeat homology <YW05>
 F:258-302/Domain: LDL receptor YWTD-containing repeat homology <YW06>
 F:319-355/Domain: LDL receptor YWTD-containing repeat homology <EG2>
 F:361-396/Domain: EGF homology <EG3>
 F:402-437/Domain: EGF homology <EG4>
 F:440-477/Domain: EGF homology <EG5>
 F:482-958/Region: EGF precursor long repeat <LR2>
 F:485-525/Domain: LDL receptor YWTD-containing repeat homology <YW07>
 F:526-568/Domain: LDL receptor YWTD-containing repeat homology <YW08>
 F:569-611/Domain: LDL receptor YWTD-containing repeat homology <YW09>
 F:612-655/Domain: LDL receptor YWTD-containing repeat homology <YW10>
 F:656-696/Domain: LDL receptor YWTD-containing repeat homology <YW11>

F:697-739/Domain: LDL receptor WTD-containing repeat homology <YW12>
 F:747-782/Domain: EGF homology <EG5>
 F:839-872/Domain: EGF homology <EG6>
 F:878-914/Domain: EGF homology <EG7>
 F:920-955/Domain: EGF homology <EG8>
 F:974-1024/Product: epidermal growth factor #status experimental <MAT>
 F:979-1015/Domain: EGF homology <EG9>
 F:1036-1060/Domain: transmembrane #status predicted <TMN>
 F:1061-1133/Domain: intracellular #status predicted <INT>
 F:342-355, 361-372, 368-381, 383-396, 402-413, 409-422, 424-437, 440-452, 448-462, 464-477, 747-75
 -1015/Disulfide bonds: #status predicted

Query Match 68.6%; Score 35; DB 1; Length 1133;
 Best Local Similarity 62.5%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
 DB 747 CLHRNGGC 754

RESULT 15

S53868
 laminin alpha-2 chain precursor - mouse
 N:Alternate names: laminin M chain; merosin heavy chain
 C:Species: Mus musculus (house mouse)
 C:Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 10-Dec-1999
 C:Accession: I49077; S50829; I48655; S31576; S53868
 R:Bernier, S.M.; Utani, A.; Sugiyama, S.; Doi, T.; Pollistina, C.; Yamada, Y.
 Matrix Biol. 14, 447-455, 1995
 A:Title: Cloning and expression of laminin alpha 2 chain (M-chain) in the mouse.
 A:Reference number: I49077; MUID:95316259; PMID:7795883
 A:Accession: I49077
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: mRNA
 A:Residues: 1-3106 <RES>
 A:Cross-references: EMBL:U12147; NID:G699109; PIDN:AAC52165.1; PID:G699110
 R:Xu, H.; Wu, X.R.; Wever, U.M.; Engvall, E.
 Nature Genet. 8, 297-302, 1994
 A:Title: Murine muscular dystrophy caused by a mutation in the laminin alpha-2 (Lama2) g
 A:Reference number: S50829; MUID:95179178; PMID:7874173
 A:Accession: S50829
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 64-281 <XUH>
 A:Cross-references: GB:S75315; NID:G833929; PIDN:AAB33573.1; PID:G833930
 R:Chang, A.C.; Wadsworth, S.; Colligan, J.E.
 J. Immunol. 151, 1789-1801, 1993
 A:Title: Expression of merosin in the thymus and its interaction with thymocytes.
 A:Reference number: I48655; MUID:93346725; PMID:8345183
 A:Accession: I48655
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: mRNA
 A:Residues: 2162-2204, 'D', 2206-2213, 'EY', 2216-2279 <RE2>
 A:Cross-references: EMBL:X69869; NID:G53055; PIDN:CAA49502.1; PID:G53056
 C:Complex: Laminins are trimers of an alpha-type, a beta-type, and a gamma-type laminin
 C:Function:
 C:Description: interact with cells and with other basement membrane proteins to promote
 C:Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-like h
 C:Keywords: basement membrane; calcium binding; coiled coil; extracellular matrix; glyco
 F:1-22/Domain: signal sequence #status predicted <SIG>
 F:23-3106/Product: laminin alpha-2 chain #status predicted <MAT>
 F:283-337/Domain: laminin-type EGF-like homology <LE01>
 F:340-407/Domain: laminin-type EGF-like homology <LE02>
 F:410-462/Domain: laminin-type EGF-like homology <LE03>
 F:465-511/Domain: laminin-type EGF-like homology <LE04>
 F:514-523/Domain: laminin-type EGF-like homology #status atypical <LE05>
 F:720-750/Domain: laminin-type EGF-like homology <LE06>
 F:753-800/Domain: laminin-type EGF-like homology <LE07>
 F:803-858/Domain: laminin-type EGF-like homology <LE08>
 F:861-911/Domain: laminin-type EGF-like homology <LE09>
 F:914-960/Domain: laminin-type EGF-like homology <LE10>
 F:963-1007/Domain: laminin-type EGF-like homology <LE11>

F:1010-1053/Domain: laminin-type EGF-like homology <LE12>
 F:1056-1099/Domain: laminin-type EGF-like homology <LE13>
 F:1102-1121/Domain: laminin-type EGF-like homology #status atypical <LE14>
 F:1123-1159/Domain: laminin-type EGF-like homology #status atypical <LE15>
 F:1162-1171/Domain: laminin-type EGF-like homology #status atypical <LE16>
 F:1376-1413/Domain: laminin-type EGF-like homology <LE18>
 F:1416-1462/Domain: laminin-type EGF-like homology <LE19>
 F:1465-1520/Domain: laminin-type EGF-like homology <LE20>
 F:1523-1567/Domain: laminin-type EGF-like homology <LE21>
 F:2166-2327/Domain: laminin G repeat homology <LG1>
 F:2360-2520/Domain: laminin G repeat homology <LG2>
 F:2546-2709/Domain: laminin G repeat homology <LG3>
 F:2785-2933/Domain: laminin G repeat homology <LG4>
 F:2960-3106/Domain: laminin G repeat homology <LG5>

Query Match 68.6%; Score 35; DB 1; Length 3106;
 Best Local Similarity 62.5%; Pred. No. 3.8e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
 DB 1012 CSHLGNCC 1019

Search completed: July 16, 2003, 13:16:15
 Job time : 11.5366 secs

THIS PAGE BLANK (USPTO)

SQL	SEQUENCE	448 AA:	49574 MW;	45203E2255D5ELD1	CRC64;
	Query Match		70.6%;	Score 36;	DB 1; Length 448;
	Best Local Similarity		85.7%;	Pred. No. 12;	
	Matches	6;	Conservative	0;	Mismatches 1; Indels 0; Gaps 0;
QY		2 LHRGNSC 8			
DB		274 LHRGTSC 280			
	RESULT 2				
VFP2_BOVIN					
ID	VFP2_BOVIN	STANDARD;	PRT;	854 AA.	
AC	O97681;				
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Vacuolar proton translocating ATPase 116 kDa subunit A isoform 2				
VE	(V-ATPase 116-kDa isoform a2).				
GN	ATP6N1B.				
OS	Bos taurus (Bovine).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OC	Bovidae; Bos.				
OX	NCBI_TaxID=9913;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99107917; PubMed=9891027;				
RA	Peng S.-B., Li X., Crider B.P., Zhou Z., Andersen P., Tsai S.J.,				
RA	Xie X.-S., Stone D.K.;				
RT	"Identification and reconstitution of an isoform of the 116-kDa				
RL	subunit of the vacuolar proton translocating ATPase.";				
RL	J. Biol. Chem. 274:2549-2555(1999).				
CC	-I- FUNCTION: PART OF THE PROTON CHANNEL OF V-ATPASES.				
CC	-I- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A				
CC	PERIPHERAL CATALYTIC V1 COMPLEX (MAIN COMPONENTS: SUBUNITS A, B,				
CC	C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE V0 PROTON PORE				
CC	COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN).				
CC	-I- SUBCELLULAR LOCATION: Integral membrane protein.				
CC	-I- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LUNG, KIDNEY, AND SPLEEN.				
CC	-I- SIMILARITY: BELONGS TO THE V-ATPASE 116 kDa SUBUNIT FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; AF1050016; AAD12058.1; "				
DR	InterPro; IPR002490; V_ATPase_sub116.				
DR	Pfam; PF01496; V_ATPase_sub_a; 1.				
KW	Hydrogen ion transport; Transmembrane; Glycoprotein.				
FT	DOMAIN 1 396				
FT	TRANSMEM 397 413				
FT	TRANSMEM 452 472				
FT	TRANSMEM 555 575				
FT	TRANSMEM 591 611				
FT	TRANSMEM 651 671				
FT	TRANSMEM 787 807				
FT	CARBOHYD 43 43				
FT	CARBOHYD 157 157				
FT	CARBOHYD 438 438				
FT	CARBOHYD 484 484				
FT	CARBOHYD 505 505				
SEQ	SEQUENCE 854 AA; 98009 MW; 8BD9A128465CED5 CRC64;				
	Query Match		70.6%;	Score 36;	DB 1; Length 854;
	Best Local Similarity		85.7%;	Pred. No. 22;	
	Matches	6;	Conservative	0;	Mismatches 1; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
 II :||:|
 Db 191 CLFQGNCC 198

RESULT 4
 DSRA_RAT STANDARD; PRT; 1175 AA.
 ID P55266;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Double-stranded RNA-specific adenosine deaminase (EC 3.5.4.-) (DRADA).
 GN ADAR OR DSRAD.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPRAIN-Wistar; TISSUE-Brain;
 RX MEDLINE=95166224; PubMed=7862132;
 RA O'Connell M.A., Krause S., Higuchi M., Hsuan J.J., Totty N.F.,
 RA Jenny A., Keller W.;
 RT "Cloning of cDNAs encoding mammalian double-stranded RNA-specific
 RT adenosine deaminase.";
 RL Mol. Cell. Biol. 15:1389-1397(1995).
 CC -1- FUNCTION: CONVERTS MULTIPLE ADENOSINES TO INOSINES AND CREATES I/U
 CC MISMATCHED BASE PAIRS IN DOUBLE-HELICAL RNA SUBSTRATES WITHOUT
 CC APPARENT SEQUENCE SPECIFICITY. DRADA HAS BEEN FOUND TO MODIFY MORE
 CC FREQUENTLY ADENOSINES IN AU-RICH REGIONS, PROBABLY DUE TO THE
 CC RELATIVE EASE OF MELTING A/U BASE PAIRS AS COMPARED TO G/C PAIRS.
 CC FUNCTIONS TO MODIFY VIRAL RNA GENOMES AND MAY BE RESPONSIBLE FOR
 CC HYPERMUTATION OF CERTAIN NEGATIVE-STRANDED VIRUSES. EDITS THE
 CC MESSENGER RNAs FOR GLUTAMATE RECEPTOR (GLUR) SUBUNITS BY SITE-
 CC SELECTIVE ADENOSINE DEAMINATION. PRODUCES LOW-LEVEL EDITING AT THE
 CC GLUR-B Q/R SITE, BUT EDITS EFFICIENTLY AT THE R/G SITE AND
 CC HOTSPOT1.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE ADENOSINE DEAMINASE EDITASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 3 DRBM (DOUBLE-STRANDED RNA-BINDING) DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 DRADA REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>).
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL: U18942; AAA65039.1; -;
 DR HSP; P25159; 1STU.
 DR InterPro: IPR002466; A_deamin.
 DR InterPro: IPR001159; DS_RBD.
 DR InterPro: IPR000607; z-alpha.
 DR Pfam: PF00035; dsxm; 3.
 DR Pfam: PF02137; A_deamin; 1.
 DR Pfam: PF02295; z-alpha; 2.
 DR SMART: SM00358; DSRM; 3.
 DR SMART: SM00550; zalpha; 2.
 DR PROSITE: PS50141; A_DEAMIN_EDITASE; 1.
 DR PROSITE: PS50139; DRADA_REPEAT; 2.
 DR PROSITE: PS50137; DS_RBD; 3.
 KW mRNA processing; Hydroxylase; Zinc; RNA-binding; Nuclear protein;
 KW Repeat.
 FT REPEAT 135 204 DRADA 1.
 FT REPEAT 243 310 DRADA 2.
 FT DOMAIN 503 519 DRBM 1.
 FT DOMAIN 614 630 DRBM 2.
 FT DOMAIN 722 738 DRBM 3.
 FT METAL 856 856 ZINC (BY SIMILARITY).

FT ACT_SITE 858 858 BY SIMILARITY.
 FT METAL 912 912 ZINC (BY SIMILARITY).
 FT METAL 982 982 ZINC (BY SIMILARITY).
 SQ SEQUENCE 1175 AA; 129911 MW; F5C7B5FFB771168C CRC64;

Query Match 69.6%; Score 35.5; DB 1; Length 1175;
 Best Local Similarity 66.7%; Pred. No. 37;
 Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 CLHR-GNSC 8
 II :||:|
 Db 572 CMHKLNSC 580

RESULT 5
 DSRA_HUMAN STANDARD; PRT; 1226 AA.
 ID P55265;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Double-stranded RNA-specific adenosine deaminase (EC 3.5.4.-) (DRADA).
 DE (136 kDa double-stranded RNA binding protein) (P136) (K88DSRBP).
 GN ADAR OR DSRAD.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=95082287; PubMed=7972084;
 RA Kim U., Wang Y., Sanford T., Zeng Y., Nishikura K.;
 RT "Molecular cloning of cDNA for double-stranded RNA adenosine
 RT deaminase, a candidate enzyme for nuclear RNA editing.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:11457-11461(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Kidney;
 RX MEDLINE=96009564; PubMed=7565688;
 RA Patterson J.B., Samuel C.E.;
 RT "Expression and regulation by interferon of a double-stranded-RNA-
 RT specific adenosine deaminase from human cells: evidence for two forms
 RT of the deaminase.";
 RL Mol. Cell. Biol. 15:5376-5388(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Placenta;
 RX MEDLINE=97172521; PubMed=9020165;
 RA Liu Y., George C.X., Patterson J.B., Samuel C.E.;
 RT "Functionally distinct double-stranded RNA-binding domains associated
 RT with alternative splice site variants of the interferon-inducible
 RT double-stranded RNA-specific adenosine deaminase.";
 RL J. Biol. Chem. 272:4419-4428(1997).
 CC -1- FUNCTION: CONVERTS MULTIPLE ADENOSINES TO INOSINES AND CREATES I/U
 CC MISMATCHED BASE PAIRS IN DOUBLE-HELICAL RNA SUBSTRATES WITHOUT
 CC APPARENT SEQUENCE SPECIFICITY. DRADA HAS BEEN FOUND TO MODIFY MORE
 CC FREQUENTLY ADENOSINES IN AU-RICH REGIONS, PROBABLY DUE TO THE
 CC RELATIVE EASE OF MELTING A/U BASE PAIRS AS COMPARED TO G/C PAIRS.
 CC FUNCTIONS TO MODIFY VIRAL RNA GENOMES AND MAY BE RESPONSIBLE FOR
 CC HYPERMUTATION OF CERTAIN NEGATIVE-STRANDED VIRUSES. EDITS THE
 CC MESSENGER RNAs FOR GLUTAMATE RECEPTOR (GLUR) SUBUNITS BY SITE-
 CC SELECTIVE ADENOSINE DEAMINATION. PRODUCES LOW-LEVEL EDITING AT THE
 CC GLUR-B Q/R SITE, BUT EDITS EFFICIENTLY AT THE R/G SITE AND
 CC HOTSPOT1.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED, HIGHEST LEVELS WERE
 CC FOUND IN BRAIN AND LUNG.
 CC -1- SIMILARITY: CONTAINS 3 DRBM (DOUBLE-STRANDED RNA-BINDING) DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 DRADA REPEATS.
 CC -1- SIMILARITY: TO DS-RNA-SPECIFIC EDITASE 1 (RED1).
 CC -1- SIMILARITY: SOME, TO THE CYTIDINE AND DEOXYCYTIDYLATE DEAMINASES
 CC FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 CC DR EMBL; U10439; AAB06697.1; -
 CC DR EMBL; U18121; AAC13782.1; -
 CC DR EMBL; U75503; AAB97118.1; JOINED.
 CC DR EMBL; U75489; AAB97118.1; JOINED.
 CC DR EMBL; U75490; AAB97118.1; JOINED.
 CC DR EMBL; U75491; AAB97118.1; JOINED.
 CC DR EMBL; U75492; AAB97118.1; JOINED.
 CC DR EMBL; U75493; AAB97118.1; JOINED.
 CC DR EMBL; U75494; AAB97118.1; JOINED.
 CC DR EMBL; U75495; AAB97118.1; JOINED.
 CC DR EMBL; U75496; AAB97118.1; JOINED.
 CC DR EMBL; U75497; AAB97118.1; JOINED.
 CC DR EMBL; U75498; AAB97118.1; JOINED.
 CC DR EMBL; U75499; AAB97118.1; JOINED.
 CC DR EMBL; U75500; AAB97118.1; JOINED.
 CC DR EMBL; U75501; AAB97118.1; JOINED.
 CC DR EMBL; U75502; AAB97118.1; JOINED.
 CC DR HSP; P19525; 1QU6.
 CC DR Genew; HGNC:225; ADAR.
 CC DR MIM; 601059; -
 CC DR InterPro; IPR002466; A_deamin.
 CC DR InterPro; IPR001159; DS_RBD.
 CC DR InterPro; IPR000607; z-alpha.
 CC DR Pfam; PF00035; dsrm; 3.
 CC DR Pfam; PF02137; A_deamin; 1.
 CC DR Pfam; PF02295; z-alpha; 2.
 CC DR SMART; SM00358; DSRM; 3.
 CC DR SMART; SM00550; Zalpha; 2.
 CC DR PROSITE; PS0141; A_DEAMIN_EDITASE; 1.
 CC DR PROSITE; PS0139; DRADA_REPEAT; 2.
 CC DR PROSITE; PS0137; DS_RBD; 3.
 CC KW mRNA processing; Hydrolase; Zinc; RNA-binding; Nuclear protein;
 KW Repeat.
 FT REPEAT 133 202 DRADA 1.
 FT REPEAT 293 360 DRADA 2.
 FT DOMAIN 553 569 DREM 1.
 FT DOMAIN 664 680 DREM 2.
 FT DOMAIN 776 792 DREM 3.
 FT METAL 910 910 ZINC (BY SIMILARITY).
 FT ACT_SITE 912 912 BY SIMILARITY.
 FT METAL 966 966 ZINC (BY SIMILARITY).
 FT METAL 1036 1036 ZINC (BY SIMILARITY).
 FT CONFLICT 384 384 K -> R (IN REF. 2 AND 3).
 SQ SEQUENCE 1226 AA; 135967 MW; 02CA60C7F881E88F CRC64;
 Query Match 69.6%; Score 35.5; DB 1; Length 1226;
 Best Local Similarity 66.7%; Pred. No. 38;
 Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 1;
 QY 1 CLHR-GNSC 8
 Db 622 CMHKLNSC 630
 RESULT 6
 CUF2_SCHPO STANDARD; PRT; 177 AA.
 ID AC 094588;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Copper-binding regulatory protein cuf2 (Copper homeostasis protein).
 GN CUF2 OR SPCC584.02.
 OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Labbe S., Thiele D.J.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RC MEDLINE=21848401; PubMed=11859360;
 RX Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Church C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., Mclellan J.,
 RA Mooney P., Moule S., Mungall K., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grynoprez B.,
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC CC -!- SIMILARITY: CONTAINS 1 COPPER-FIST DNA-BINDING DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ243834; CAB52306.1; -
 CC EMBL; AL032824; CAA21819.1; -
 CC InterPro; IPR001083; Copper-fist.
 CC Pfam; PF00649; Copper-fist; 1.
 CC PRINTS; PR00617; COPPERFIST.
 CC ProDom; PD009050; Copper-fist; 1.
 CC SMART; SM00412; Cu_FIST; 1.
 CC PROSITE; PS01119; COPPER_FIST_1; FALSE_NEG.
 CC PROSITE; PS00073; COPPER_FIST_2; 1.
 CC Transcription regulation; DNA-binding; Copper; Nuclear protein.
 KW DNA_BIND 1 40 COPPER-FIST.
 FT METAL 11 11 ZINC (BY SIMILARITY).
 FT METAL 14 14 ZINC (BY SIMILARITY).
 FT METAL 23 23 ZINC (BY SIMILARITY).
 FT METAL 25 25 ZINC (BY SIMILARITY).
 SQ SEQUENCE 177 AA; 20070 MW; CFD0599D376F283 CRC64;
 Query Match 68.6%; Score 35; DB 1; Length 177;
 Best Local Similarity 70.0%; Pred. No. 7.4;
 Matches 7; Conservative 1; Mismatches 0; Indels 2; Gaps 1;
 QY 1 CL-HRNSC 8
 Db 14 CLRGHGSSC 23

RESULT 7

NP4_RHOPR STANDARD; PRT; 205 AA.
 AC Q94734;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Nitrophorin 4 precursor (NP4).
 OS Rhodnius prolixus (Tritatomid bug).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
 OC Heteroptera; Panheteroptera; Cimicomorpha; Reduviidae; Triatominae;
 OC Rhodnius.
 OX NCBI_TaxID=13249;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RA Champagne D.E., Ribeiro J.M.C.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 22-205.
 RX MEDLINE=9845513; PubMed=9782054;
 RA Andersen J.F., Weichsel A., Balfour C.A., Champagne D.E.,
 RA Montfort W.R.;
 RT "The crystal structure of nitrophorin 4 at 1.5-A resolution: transport
 of nitric oxide by a lipocalin-based heme protein.";
 RL Structure 6:1315-1327(1998).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS).
 RX MEDLINE=2033693; PubMed=10876239;
 RA Weichsel A., Andersen J.F., Roberts S.A., Montfort W.R.;
 RT "Nitric oxide binding to nitrophorin 4 induces complete distal pocket
 burial.";
 RL Nat. Struct. Biol. 7:551-554(2000).
 CC -1- FUNCTION: HEME-BASED PROTEIN THAT DELIVER NITRIC OXIDE GAS (NO) TO
 THE VICTIM WHILE FEEDING, RESULTING IN VASODILATION AND INHIBITION
 OF PLATELET AGGREGATION. ALSO BIND TIGHTLY TO HISTAMINE, WHICH IS
 RELEASED BY THE HOST TO INDUCE WOUND HEALING (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: EXPRESSED BY THE SALIVARY GLANDS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U70584; AAB09590.1;
 DR NP4; LNP4; 05-AUG-98.
 DR PDB; LEQD; 03-MAY-00.
 DR PDB; LERX; 03-MAY-00.
 DR PDB; LD3S; 07-JUL-00.
 DR InterPro; IPR002351; Nitrophorin.
 DR Pfam; PF02087; Nitrophorin; 1.
 KW Heme; Vasodilator; Signal; 3D-structure.
 FT SIGNAL 1 21
 FT CHAIN 22 205 NITROPHORIN 4.
 FT DISULFID 23 143
 FT DISULFID 62 192
 FT METAL 80 80 IRON (HEME PROXIMAL LIGAND).
 SQ SEQUENCE 205 AA; 22406 MW; 9BACAI2001527F2 CRC64;
 Query Match 68.6%; Score 35; DB 1; Length 205;
 Best Local Similarity 83.3%; Pred. No. 8.5;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CLHGRN 6
 DB 143 CLHKN 148

RESULT 8

NP1_RHOPR STANDARD; PRT; 207 AA.
 AC Q26239;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Nitrophorin 1 precursor (NP1).
 OS Rhodnius prolixus (Tritatomid bug).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
 OC Heteroptera; Panheteroptera; Cimicomorpha; Reduviidae; Triatominae;
 OC Rhodnius.
 OX NCBI_TaxID=13249;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Salivary gland;
 RX MEDLINE=95238361; PubMed=7721773;
 RA Champagne D.E., Nussenzweig R.H., Ribeiro J.M.C.;
 RT "Purification, partial characterization, and cloning of nitric oxide-
 carrying heme proteins (nitrophorins) from salivary glands of the
 blood-sucking insect Rhodnius prolixus.";
 RL J. Biol. Chem. 270:8691-8695(1995).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RC TISSUE=Salivary gland;
 RX MEDLINE=98206297; PubMed=9546222;
 RA Weichsel A., Andersen J.F., Champagne D.E., Walker F.A.,
 RA Montfort W.R.;
 RT "Crystal structures of a nitric oxide transport protein from a blood-
 sucking insect.";
 RL Nat. Struct. Biol. 5:304-309(1998).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RA Ding X.D., Weichsel A., Andersen J.F., Shokhireva T.K., Balfour C.,
 RA Plerik A., Averill B.A., Montfort W.R., Walker F.A.;
 RL Submitted (JUN-1998) to the PDB data bank.
 CC -1- FUNCTION: HEME-BASED PROTEIN THAT DELIVER NITRIC OXIDE GAS (NO) TO
 THE VICTIM WHILE FEEDING, RESULTING IN VASODILATION AND INHIBITION
 OF PLATELET AGGREGATION. ALSO BIND TIGHTLY TO HISTAMINE, WHICH IS
 RELEASED BY THE HOST TO INDUCE WOUND HEALING.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: EXPRESSED BY THE SALIVARY GLANDS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L39654; AAA74407.1;
 DR PDB; 1NP1; 27-MAY-98.
 DR PDB; 2NP1; 27-MAY-98.
 DR PDB; 3NP1; 12-AUG-98.
 DR PDB; 4NP1; 11-NOV-98.
 DR InterPro; IPR002351; Nitrophorin.
 DR Pfam; PF02087; Nitrophorin; 1.
 KW Heme; Vasodilator; Signal; 3D-structure.
 FT SIGNAL 1 23
 FT CHAIN 24 207 NITROPHORIN 1.
 FT DISULFID 25 145
 FT DISULFID 64 194
 FT METAL 82 82 IRON (HEME PROXIMAL LIGAND).
 SQ SEQUENCE 207 AA; 22763 MW; 4AB9EE803FDA0EB8 CRC64;
 Query Match 68.6%; Score 35; DB 1; Length 207;
 Best Local Similarity 83.3%; Pred. No. 8.6;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

AC Q60675; Q05003; Q64061;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Laminin alpha-2 chain precursor (Laminin M chain) (Merosin heavy
chain).
DE LAMA2.
GN Mus musculus (Mouse).
OS Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-FVB/N; TISSUE=Embryo, and Heart;
RX MEDLINE=95316259; PubMed=7795883;
RA Bernier S.M., Utani A., Sugiyama S., Doi T., Pollistina C.,
RA Yamada Y.;
RT "Cloning and expression of laminin alpha 2 chain (M-chain) in the
mouse.";
RL Matrix Biol. 14:447-455(1995).
RN [2]
RN SEQUENCE OF 2162-2279 FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Thymus;
RX MEDLINE=93346725; PubMed=8345183;
RA Chang A.C., Wadsworth S., Colligan J.E.;
RT "Expression of merosin in the thymus and its interaction with
thymocytes.";
RL J. Immunol. 151:1789-1801(1993).
RN [3]
RN SEQUENCE OF 64-281 FROM N.A.
RX MEDLINE=95179178; PubMed=7874173;
RA Xu H., Wu X.R., Wewer U.M., Engvall E.;
RT "Murine muscular dystrophy caused by a mutation in the laminin alpha
2 (Lama2) gene.";
RL Nat. Genet. 8:297-302(1994).
RN [4]
RN X-RAY CRYSTALLOGRAPHY (2.35 ANGSTROMS) OF 2932-3106.
RX MEDLINE=20085745; PubMed=10619025;
RA Hohenester E., Tisi D., Talts J.F., Timpl R.;
RT "The crystal structure of a laminin G-like module reveals the
molecular basis of alpha-dystroglycan binding to laminins, perlecan,
and agrin.";
RL Mol. Cell 4:783-792(1999).
CC -1- FUNCTION: Binding to cells via a high affinity receptor, laminin
is thought to mediate the attachment, migration, and organization
of cells into tissues during embryonic development by interacting
with other extracellular matrix components.
CC -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
different polypeptide chains (alpha, beta, gamma), which are bound
to each other by disulfide bonds into a cross-shaped molecule
comprising one long and three short arms with globules at each
end.
CC THE ALPHA-2 CHAIN IS A SUBUNIT OF LAMININ-2 (MEROSIN) AND LAMININ-
4 (S-MEROSIN).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
COMPONENT).
CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -1- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.
CC -1- DISEASE: DEFECTS IN LAMA2 ARE A CAUSE OF MURINE MUSCULAR DYSTROPHY
(DY2J).
CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -1- SIMILARITY: CONTAINS 17 LAMININ EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 LAMININ DOMAINS IV.
CC -1- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
CC -----
DR EMBL; UI2147; AAC52165.1; -
DR EMBL; X69869; CAA49502.1; -
DR EMBL; S75315; AAB33573.1; -
DR PDB; 1QU0; O3-DEC-99.
DR MGD; MGI:99912; Lama2.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001886; LamNT.
DR InterPro; IPR000034; Laminin_B.
DR InterPro; IPR002049; Laminin_G.
DR InterPro; IPR001791; Laminin_G.
DR Pfam; PF00052; laminin_B; 2.
DR Pfam; PF00053; laminin_EGF; 15.
DR Pfam; PF00054; laminin_G; 5.
DR Pfam; PF00055; laminin_Nterm; 1.
DR PRINTS; PR00011; EGF_LAMININ.
DR ProDom; PD002082; LamNT; 1.
DR ProDom; PD003031; Laminin_B; 2.
DR SMART; SM00180; EGF_Lam; 15.
DR SMART; SM00001; EGF_Like; 1.
DR SMART; SM00281; Lamb; 2.
DR SMART; SM00282; LambG; 5.
DR SMART; SM00136; LamNT; 1.
DR PROSITE; PS00022; EGF_1; 11.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 14.
DR PROSITE; PS00025; LAM_G_DOMAIN; 5.
KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
Laminin EGF-like domain; Cell adhesion; Repeat; Signal; 3D-structure.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 3106 LAMININ ALPHA-2 CHAIN.
FT DOMAIN 23 282 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 283 339 LAMININ EGF-LIKE 1.
FT DOMAIN 340 409 LAMININ EGF-LIKE 2.
FT DOMAIN 410 464 LAMININ EGF-LIKE 3.
FT DOMAIN 465 513 LAMININ EGF-LIKE 4.
FT DOMAIN 514 523 LAMININ EGF-LIKE 5 (N-TERMINAL).
FT DOMAIN 524 719 LAMININ DOMAIN IV 1 (DOMAIN IV B).
FT DOMAIN 720 752 LAMININ EGF-LIKE 5 (C-TERMINAL).
FT DOMAIN 753 802 LAMININ EGF-LIKE 6.
FT DOMAIN 803 860 LAMININ EGF-LIKE 7.
FT DOMAIN 861 913 LAMININ EGF-LIKE 8.
FT DOMAIN 914 962 LAMININ EGF-LIKE 9.
FT DOMAIN 963 1009 LAMININ EGF-LIKE 10.
FT DOMAIN 1010 1055 LAMININ EGF-LIKE 11.
FT DOMAIN 1056 1101 LAMININ EGF-LIKE 12.
FT DOMAIN 1102 1161 LAMININ EGF-LIKE 13.
FT DOMAIN 1162 1171 LAMININ EGF-LIKE 14 (N-TERMINAL).
FT DOMAIN 1172 1375 LAMININ DOMAIN IV 2 (DOMAIN IV A).
FT DOMAIN 1376 1415 LAMININ EGF-LIKE 14 (C-TERMINAL).
FT DOMAIN 1416 1464 LAMININ EGF-LIKE 15.
FT DOMAIN 1465 1522 LAMININ EGF-LIKE 16.
FT DOMAIN 1523 1569 LAMININ EGF-LIKE 17.
FT DOMAIN 1570 2140 DOMAIN II AND I.
FT DOMAIN 2141 2324 LAMININ G-LIKE 1.
FT DOMAIN 2326 2517 LAMININ G-LIKE 2.
FT DOMAIN 2522 2706 LAMININ G-LIKE 3.
FT DOMAIN 2759 2930 LAMININ G-LIKE 4.
FT DOMAIN 2929 3106 LAMININ G-LIKE 5.
FT DOMAIN 1662 1863 COILED COIL (POTENTIAL).
FT DOMAIN 1923 2146 COILED COIL (POTENTIAL).
FT DISULFID 283 292 BY SIMILARITY.
FT DISULFID 285 303 BY SIMILARITY.
FT DISULFID 305 314 BY SIMILARITY.
FT DISULFID 317 337 BY SIMILARITY.
FT DISULFID 340 349 BY SIMILARITY.
FT DISULFID 342 374 BY SIMILARITY.
FT DISULFID 377 386 BY SIMILARITY.
FT DISULFID 389 407 BY SIMILARITY.
FT DISULFID 410 422 BY SIMILARITY.
FT DISULFID 412 438 BY SIMILARITY.
FT DISULFID 440 449 BY SIMILARITY.

FT DISULFID 452 462 BY SIMILARITY.
 FT DISULFID 465 478 BY SIMILARITY.
 FT DISULFID 467 482 BY SIMILARITY.
 FT DISULFID 484 493 BY SIMILARITY.
 FT DISULFID 496 511 BY SIMILARITY.
 FT DISULFID 753 762 BY SIMILARITY.
 FT DISULFID 755 769 BY SIMILARITY.
 FT DISULFID 772 781 BY SIMILARITY.
 FT DISULFID 784 800 BY SIMILARITY.
 FT DISULFID 803 818 BY SIMILARITY.
 FT DISULFID 805 828 BY SIMILARITY.
 FT DISULFID 831 840 BY SIMILARITY.
 FT DISULFID 843 858 BY SIMILARITY.
 FT DISULFID 861 875 BY SIMILARITY.
 FT DISULFID 863 882 BY SIMILARITY.
 FT DISULFID 885 894 BY SIMILARITY.
 FT DISULFID 897 911 BY SIMILARITY.
 FT DISULFID 914 926 BY SIMILARITY.
 FT DISULFID 916 933 BY SIMILARITY.
 FT DISULFID 935 944 BY SIMILARITY.
 FT DISULFID 947 960 BY SIMILARITY.
 FT DISULFID 963 975 BY SIMILARITY.
 FT DISULFID 965 981 BY SIMILARITY.
 FT DISULFID 983 992 BY SIMILARITY.
 FT DISULFID 993 1007 BY SIMILARITY.
 FT DISULFID 1010 1019 BY SIMILARITY.
 FT DISULFID 1012 1026 BY SIMILARITY.
 FT DISULFID 1028 1037 BY SIMILARITY.
 FT DISULFID 1040 1053 BY SIMILARITY.
 FT DISULFID 1056 1068 BY SIMILARITY.
 FT DISULFID 1058 1075 BY SIMILARITY.
 FT DISULFID 1077 1086 BY SIMILARITY.
 FT DISULFID 1089 1099 BY SIMILARITY.
 FT DISULFID 1416 1425 BY SIMILARITY.
 FT DISULFID 1418 1432 BY SIMILARITY.
 FT DISULFID 1435 1444 BY SIMILARITY.
 FT DISULFID 1447 1462 BY SIMILARITY.
 FT DISULFID 1465 1480 BY SIMILARITY.
 FT DISULFID 1467 1490 BY SIMILARITY.
 FT DISULFID 1493 1502 BY SIMILARITY.
 FT DISULFID 1505 1520 BY SIMILARITY.
 FT DISULFID 1523 1535 BY SIMILARITY.
 FT DISULFID 1525 1542 BY SIMILARITY.
 FT DISULFID 1544 1553 BY SIMILARITY.
 FT DISULFID 1556 1567 BY SIMILARITY.
 FT DISULFID 1570 1570 INTERCHAIN (PROBABLE).
 FT DISULFID 1574 1574 INTERCHAIN (PROBABLE).
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 359 359 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
 Query Match 68.6%; Score 35; DB 1; Length 3106;
 Best Local Similarity 62.5%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 CLHRGNSC 8
 Db 1012 CSHLGNCC 1019
 RESULT 11
 LMA2 HUMAN
 ID LMA2 HUMAN STANDARD; PRT: 3110 AA.
 AC P24043; Q14736; Q93022;
 DT 01-WAR-1992 (Rel. 21, created)
 DT 01-NOV-1997 (Rel. 35, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Laminin alpha-2 chain precursor (laminin M chain) (Merosin heavy chain).
 GN LMA2 OR LAMM.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=94124633; PubMed=8294519;
 RA Vuolteenaho R., Nissinen M., Sainio K., Byers M., Eddy R.,
 RA Vuolteenaho R., Shows T.B., Sariola H., Engvall E., Tryggvason K.;
 RA "Human laminin M chain (merosin): complete primary structure,
 RT chromosomal assignment, and expression of the M and A chain in human
 RT fetal tissues."; J. Cell Biol. 124:381-394(1994).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=97066955; PubMed=8910357;
 RX Zhang X., Vuolteenaho R., Tryggvason K.;
 RA "Structure of the human laminin alpha2-chain gene (LAMA2), which is
 RT affected in congenital muscular dystrophy."; J. Biol. Chem. 271:27664-27669(1996).
 RL [3]
 RN SEQUENCE OF 1981-3110 FROM N.A., AND PARTIAL SEQUENCE.
 RP TISSUE=Placenta;
 RC MEDLINE=90238994; PubMed=2185464;
 RX Ehrig K., Leivo I., Agravas W.S., Ruoslahti E., Engvall E.;
 RA "Merosin, a tissue-specific basement membrane protein, is a
 RT laminin-like protein."; Proc. Natl. Acad. Sci. U.S.A. 87:3264-3268(1990).
 RL [4]
 RN VARIANTS GLN-545; HIS-619; LEU-919; HIS-2586 AND LYS-2614.
 RP Panicker S.G., Mendell J.T., Chen L., Feng B., Sahenk Z.,
 RA Marzluf G.A., Amato A.A., Mendell J.R.;
 RA "Novel single base polymorphisms and rare sequence variants in
 RT the laminin 2-chain coding region detected by RNA/SSCP analysis."; Hum. Mutat. 13:174-174(1999).
 RL [5]
 RN ERRATUM.
 RP Panicker S.G., Mendell J.T., Chen L., Feng B., Sahenk Z.,
 RA Marzluf G.A., Amato A.A., Mendell J.R.;
 RL Hum. Mutat. 13:340-340(1999).
 CC -!- FUNCTION: Binding to cells via a high affinity receptor, laminin
 CC is thought to mediate the attachment, migration, and organization
 CC of cells into tissues during embryonic development by interacting
 CC with other extracellular matrix components.
 CC -!- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
 CC different polypeptide chains (alpha, beta, gamma), which are bound
 CC to each other by disulfide bonds into a cross-shaped molecule
 CC comprising one long and three short arms with globules at each
 CC end.
 CC THE ALPHA-2 CHAIN IS A SUBUNIT OF LAMININ-2 (MEROSIN) AND LAMININ-
 CC 4 (S-MEROSIN).
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
 CC MEMBRANES (MAJOR COMPONENT).
 CC -!- TISSUE SPECIFICITY: PLACENTA, STRIATED MUSCLE, PERIPHERAL NERVE,
 CC CARDIAC MUSCLE, PANCREAS, LUNG, SPLEEN, KIDNEY, ADRENAL GLAND,
 CC SKIN, TESTIS, MENINGES, CHOROID PLEXUS, AND SOME OTHER REGIONS OF
 CC THE BRAIN; NOT IN LIVER, THYMUS AND BONE.
 CC -!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
 CC -!- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.
 CC -!- DISEASE: DEFECTS IN LAMA2 ARE THE CAUSE OF MEROSIN-DEFICIENT
 CC CONGENITAL MUSCULAR DYSTROPHY (MCD).
 CC -!- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
 CC -!- SIMILARITY: CONTAINS 17 LAMININ EGF-LIKE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 2 LAMININ DOMAINS IV.
 CC -!- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

RT "Common carp metallothionein gene.";
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS (BY SIMILARITY).
 CC -1- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:
 CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
 CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11
 CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
 CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.
 CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF002161; AAB61577.1; -;
 CC EMBL; AF002162; AAB61578.1; -;
 CC EMBL; AF001983; AAB70467.1; -;
 CC HSSP; P02795; 1MHU.
 CC InterPro: IPR0003019; Metallthion.
 CC InterPro: IPR000006; Metallthion_vert.
 CC Pfam; PF00131; metalthio; 1.
 CC PRINTS; PR00860; MTVERTEBRATE.
 CC PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
 CC Metal-binding; Metal-thiolate cluster.
 KW DOMAIN 1 28 BETA.
 FT DOMAIN 29 60 ALPHA.
 FT METAL 4 4 CLUSTER B.
 FT METAL 6 6 CLUSTER B.
 FT METAL 12 12 CLUSTER B.
 FT METAL 14 14 CLUSTER B.
 FT METAL 18 18 CLUSTER B.
 FT METAL 20 20 CLUSTER B.
 FT METAL 23 23 CLUSTER B.
 FT METAL 25 25 CLUSTER B.
 FT METAL 28 28 CLUSTER A.
 FT METAL 32 32 CLUSTER A.
 FT METAL 33 33 CLUSTER A.
 FT METAL 35 35 CLUSTER A.
 FT METAL 36 36 CLUSTER A.
 FT METAL 40 40 CLUSTER A.
 FT METAL 43 43 CLUSTER A.
 FT METAL 47 47 CLUSTER A.
 FT METAL 49 49 CLUSTER A.
 FT METAL 54 54 CLUSTER A.
 FT METAL 58 58 CLUSTER A.
 FT METAL 59 59 CLUSTER A.
 SQ SEQUENCE 60 AA; 5999 MW; 54389FF53F1BB701 CRC64;

Query Match 66.78; Score 34; DB 1; Length 60;
 Best Local Similarity 62.5%; Pred. No. 4;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
 Db 47 CVCKGNSC 54
 1: |||||

RESULT 13
 ID MT2_CYPCA STANDARD; PRT; 60 AA.
 AC Q91910;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metallothionein II (MT-II).
 GN MT2.
 OS Cyprinus carpio (Common carp).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Cyprinus.
 OX NCBI_TaxID=7962;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hermesz E., Abraham M., Nemcsok J.;
 RT "Differential expression of two metallothionein genes in common
 RT carp.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS (BY SIMILARITY).
 CC -1- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:
 CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
 CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11
 CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
 CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.
 CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF249875; AAF64232.1; -;
 CC HSSP; P02795; 1MHU.
 CC InterPro: IPR003019; Metallthion.
 CC InterPro: IPR000006; Metallthion_vert.
 CC Pfam; PF00131; metalthio; 1.
 CC PRINTS; PR00860; MTVERTEBRATE.
 CC PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
 CC Metal-binding; Metal-thiolate cluster.
 KW DOMAIN 1 28 BETA.
 FT DOMAIN 29 60 ALPHA.
 FT METAL 4 4 CLUSTER B (BY SIMILARITY).
 FT METAL 6 6 CLUSTER B (BY SIMILARITY).
 FT METAL 12 12 CLUSTER B (BY SIMILARITY).
 FT METAL 14 14 CLUSTER B (BY SIMILARITY).
 FT METAL 18 18 CLUSTER B (BY SIMILARITY).
 FT METAL 20 20 CLUSTER B (BY SIMILARITY).
 FT METAL 23 23 CLUSTER B (BY SIMILARITY).
 FT METAL 25 25 CLUSTER B (BY SIMILARITY).
 FT METAL 28 28 CLUSTER B (BY SIMILARITY).
 FT METAL 32 32 CLUSTER A (BY SIMILARITY).
 FT METAL 33 33 CLUSTER A (BY SIMILARITY).
 FT METAL 35 35 CLUSTER A (BY SIMILARITY).
 FT METAL 36 36 CLUSTER A (BY SIMILARITY).
 FT METAL 40 40 CLUSTER A (BY SIMILARITY).
 FT METAL 43 43 CLUSTER A (BY SIMILARITY).
 FT METAL 47 47 CLUSTER A (BY SIMILARITY).
 FT METAL 49 49 CLUSTER A (BY SIMILARITY).
 FT METAL 54 54 CLUSTER A (BY SIMILARITY).
 FT METAL 58 58 CLUSTER A (BY SIMILARITY).
 FT METAL 59 59 CLUSTER A (BY SIMILARITY).
 SQ SEQUENCE 60 AA; 5962 MW; 54389F7D378BB701 CRC64;

Query Match 66.78; Score 34; DB 1; Length 60;
 Best Local Similarity 62.5%; Pred. No. 4;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
 Db 47 CVCKGNSC 54
 1: |||||

RESULT 14
 ID MTA_CYP5I STANDARD; PRT; 60 AA.
 AC Q92044;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metallothionein A (MT A).
 GN MTA.
 OS Cyprinodon sp. (Pupfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
 OC Cyprinodontiformes; Cyprinodontidae; Cyprinodon.
 OX NCBI_TaxID=48417;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Kille P., Olsson P.E.;
 RT "The use of metallothionein genes for determining the phylogenetic
 and evolutionary relationship between extant teleosts";
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS (BY SIMILARITY).
 CC -1- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:
 CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
 CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11
 CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
 CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.
 CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X97273; CAA65928.1; -
 DR HSSP; P28184; LJI9
 DR InterPro; IPR003019; Metallthion.
 DR Pfam; PF00131; metalthio; 1.
 DR PRINTS; PR00860; MTVERTEBRATE.
 DR PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
 KW Metal-binding; Metal-thiolate cluster; Zinc.
 FT DOMAIN 1 28 BETA.
 FT DOMAIN 29 60 ALPHA.
 FT METAL 4 4 CLUSTER B (BY SIMILARITY).
 FT METAL 6 6 CLUSTER B (BY SIMILARITY).
 FT METAL 12 12 CLUSTER B (BY SIMILARITY).
 FT METAL 14 14 CLUSTER B (BY SIMILARITY).
 FT METAL 18 18 CLUSTER B (BY SIMILARITY).
 FT METAL 20 20 CLUSTER B (BY SIMILARITY).
 FT METAL 23 23 CLUSTER B (BY SIMILARITY).
 FT METAL 25 25 CLUSTER B (BY SIMILARITY).
 FT METAL 28 28 CLUSTER B (BY SIMILARITY).
 FT METAL 32 32 CLUSTER A (BY SIMILARITY).
 FT METAL 33 33 CLUSTER A (BY SIMILARITY).
 FT METAL 35 35 CLUSTER A (BY SIMILARITY).
 FT METAL 36 36 CLUSTER A (BY SIMILARITY).
 FT METAL 40 40 CLUSTER A (BY SIMILARITY).
 FT METAL 43 43 CLUSTER A (BY SIMILARITY).
 FT METAL 47 47 CLUSTER A (BY SIMILARITY).
 FT METAL 49 49 CLUSTER A (BY SIMILARITY).
 FT METAL 54 54 CLUSTER A (BY SIMILARITY).
 FT METAL 58 58 CLUSTER A (BY SIMILARITY).
 FT METAL 59 59 CLUSTER A (BY SIMILARITY).
 SQ SEQUENCE 60 AA; 6095 MW; F73A5EE39CDBD971 CRC64;
 Query Match 66.7%; Score 34; DB 1; Length 60;
 Best Local Similarity 62.5%; Pred. No. 4;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 OY 1 CLHRGNSC 8
 Db |:::|||||
 47 CVCKGNSC 54

RESULT 15
 MT_BRARE
 ID MT_BRARE STANDARD; PRT; 60 AA.
 AC P52722;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metallothionein (MT).
 GN MT.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Kille P., Olsson P.E.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS (BY SIMILARITY).
 CC -1- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:
 CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
 CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11
 CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
 CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.
 CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X97278; CAA65933.1; -
 DR HSSP; P02795; LMHU
 DR ZFIN; ZDB-GENE-990415-159; mt.
 DR InterPro; IPR003019; Metallthion.
 DR Pfam; PF00131; metalthio; 1.
 DR PRINTS; PR00860; MTVERTEBRATE.
 DR PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
 KW Metal-binding; Metal-thiolate cluster.
 FT DOMAIN 1 28 BETA.
 FT DOMAIN 29 60 ALPHA.
 FT METAL 4 4 CLUSTER B.
 FT METAL 6 6 CLUSTER B.
 FT METAL 12 12 CLUSTER B.
 FT METAL 14 14 CLUSTER B.
 FT METAL 18 18 CLUSTER B.
 FT METAL 20 20 CLUSTER B.
 FT METAL 23 23 CLUSTER B.
 FT METAL 25 25 CLUSTER B.
 FT METAL 28 28 CLUSTER B.
 FT METAL 32 32 CLUSTER A.
 FT METAL 33 33 CLUSTER A.
 FT METAL 35 35 CLUSTER A.
 FT METAL 36 36 CLUSTER A.
 FT METAL 40 40 CLUSTER A.
 FT METAL 43 43 CLUSTER A.
 FT METAL 47 47 CLUSTER A.
 FT METAL 49 49 CLUSTER A.
 FT METAL 54 54 CLUSTER A.
 FT METAL 58 58 CLUSTER A.
 FT METAL 59 59 CLUSTER A.
 SQ SEQUENCE 60 AA; 5960 MW; E5223E7C292BB700 CRC64;
 Query Match 66.7%; Score 34; DB 1; Length 60;
 Best Local Similarity 62.5%; Pred. No. 4;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
I: :||||
Db 47 CVCKGNSC 54

Search completed: July 16, 2003, 13:13:15
Job time : 7.46341 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: July 16, 2003, 13:07:18 ; Search time 21.6585 seconds
(without alignments)
76.107 Million cell updates/sec

Title: US-09-910-582B-9
Perfect score: 51
Sequence: 1 CLHRGNSC 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*
1: sp.archaea.*
2: sp.bacteria.*
3: sp.fungi.*
4: sp.human.*
5: sp.invertebrate.*
6: sp.mammal.*
7: sp.mhc.*
8: sp.organelle.*
9: sp.phage.*
10: sp.plant.*
11: sp.rodent.*
12: sp.virus.*
13: sp.vertebrate.*
14: sp.unclassified.*
15: sp.rvirus.*
16: sp.bacteriap.*
17: sp.archeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	76.5	268	11 Q9DCQ5	Q9dcq5 mus musculus
2	39	76.5	347	11 Q8VE03	Q8ve03 mus musculus
3	39	76.5	631	10 Q9SRZ8	Q9srz8 arabidopsis
4	38	74.5	923	13 Q90ZN8	Q90zn8 brachydanio
5	38	74.5	1743	5 Q9XWX5	Q9xwx5 caenorhabdi
6	37	72.5	452	16 Q97H85	Q97h85 clostridium
7	36.5	71.6	524	3 Q74832	Q74832 schizosacch
8	36	70.6	106	8 P92546	P92546 arabidopsis
9	36	70.6	106	10 P93299	P93299 arabidopsis
10	36	70.6	179	17 Q8TTL2	Q8ttl2 methanosarc
11	36	70.6	327	4 Q8TC85	Q8tc85 homo sapien
12	36	70.6	1396	5 Q9NEC5	Q9nec5 leishmania
13	36	70.6	1731	4 Q8WY30	Q8wy30 homo sapien
14	36	70.6	4599	4 Q9NZR2	Q9nzz2 mus musculus
15	36	70.6	4599	11 Q9JTI8	Q9jti8 mus musculus
16	35.5	69.6	634	11 Q99MU4	Q99mu4 mus musculus

17	35.5	69.6	660	11 Q99MU2	Q99mu2 mus musculus
18	35.5	69.6	1152	11 Q70375	Q70375 mus musculus
19	35.5	69.6	1152	11 Q99MU7	Q99mu7 mus musculus
20	35.5	69.6	1178	11 Q99MU3	Q99mu3 mus musculus
21	35.5	69.6	1181	4 Q43859	Q43859 homo sapien
22	35.5	69.6	1200	4 Q43860	Q43860 homo sapien
23	35.5	69.6	1221	4 Q9BYM4	Q9bym4 homo sapien
24	35.5	69.6	1226	4 Q15223	Q15223 homo sapien
25	35	68.6	192	5 Q18330	Q18330 caenorhabdi
26	35	68.6	222	16 Q9CLB4	Q9cle4 pasteurella
27	35	68.6	418	5 Q21941	Q21941 caenorhabdi
28	35	68.6	423	5 Q9XVP8	Q9xvp8 caenorhabdi
29	35	68.6	440	5 Q93646	Q93646 caenorhabdi
30	35	68.6	1401	10 Q94LJ6	Q94lj6 oryza sativ
31	35	68.6	1905	5 Q9V4V8	Q9v4v8 drosophila
32	34	66.7	38	7 Q31250	Q31250 peromyscus
33	34	66.7	49	13 Q98S15	Q98s15 leuciscus c
34	34	66.7	82	3 Q02039	Q02039 rhynchospo
35	34	66.7	124	11 Q9CVM5	Q9cvw5 mus musculus
36	34	66.7	196	5 Q24346	Q24346 drosophila
37	34	66.7	198	4 Q9NPS2	Q9nps2 homo sapien
38	34	66.7	229	10 Q8RZ37	Q8rz37 oryza sativ
39	34	66.7	290	10 Q82555	Q82555 blastocysti
40	34	66.7	290	10 Q82556	Q82556 blastocysti
41	34	66.7	290	10 Q82557	Q82557 blastocysti
42	34	66.7	290	10 Q82558	Q82558 blastocysti
43	34	66.7	290	10 Q9M504	Q9m504 blastocysti
44	34	66.7	290	10 Q82559	Q82559 blastocysti
45	34	66.7	299	17 Q30200	Q30200 archaeoglob

ALIGNMENTS

RESULT 1

ID	Q9DCQ5	PRELIMINARY:	PRT:	268 AA.
AC	Q9DCQ5:			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DE	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	4930453N24RIK protein.			
GN	4930453N24RIK.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=KIDNEY;			
EX	MEDLINE=21085660; Pubmed=11217851;			
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA	Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,			
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,			
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,			
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,			
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,			
RA	Hayashizaki Y.;			
RT	"Functional annotation of a full-length mouse cDNA collection."			
RL	Nature 409:685-690(2001).			
DR	EMBL; AK002578; BAB22202.1;			
DR	MGD; MGI:1914859; 4930453N24RIK.			

DR InterPro: IPR000169; SHprot.acsite.
DR PROSITE: PS00639; THIOI_PROTEASE_HIS; UNKNOWN_1.
SQ SEQUENCE 268 AA; 30644 MW; 4DB4C485ECCA40DC CRC64;

Query Match 76.5%; Score 39; DB 11; Length 268;
Best Local Similarity 85.7%; Pred. No. 7;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LHRGNSC 8
Db 237 VHRGNSC 243
:|||||

RESULT 2
Q8VE03 PRELIMINARY; PRT; 347 AA.
ID Q8VE03
AC Q8VE03; 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 21, Last sequence update)
DE Similar to RIKEN CDNA 4930453N24 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020029; AAH20029.1;
DR InterPro: IPR000169; SHprot.acsite.
DR PROSITE; PS00639; THIOI_PROTEASE_HIS; UNKNOWN_1.
SQ SEQUENCE 347 AA; 39373 MW; 41D7C9A2CC284D38 CRC64;

Query Match 76.5%; Score 39; DB 11; Length 347;
Best Local Similarity 85.7%; Pred. No. 8;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LHRGNSC 8
Db 236 VHRGNSC 242
:|||||

RESULT 3
Q9SRZ8 PRELIMINARY; PRT; 631 AA.
ID Q9SRZ8
AC Q9SRZ8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE Fl2p19.8 protein.
GN Fl2p19.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosidis II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S., Lee J.M., Li J., E.,
RA Gonzalez A., Liu A., Sakano H., Vaysberg M., Chin C., Choi E.,
RA Chiu J., Altati H., Araujo R., Brooks S., Buehler E., Chao Q.,
RA Conn L., Conway A., Dunn P., Hansen N., Howng B., Huizar L.,
RA Johnson-Hopson C., Khan S., Kim C., Lam B., Nguyen M., Palm C.,
RA Rowley D., Shinn P., Southwick A., Tambunga G., Walker M., Davis R.W.,
RA Ecker J.R., Federspiel N.A., Theologis A.;
RA "Arabidopsis thaliana chromosome 1 BAC Fl2p19 sequence."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC009513; AAF06052.1;
DR InterPro: IPR003441; NAM.
DR Pfam; PF02365; NAM; 1.
SQ SEQUENCE 631 AA; 71907 MW; F07E90365B1EDE4F CRC64;

Query Match 76.5%; Score 39; DB 10; Length 631;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CLHRGNS 7
Db 547 CVHRGNS 553
:|||||

RESULT 4
Q90ZNS PRELIMINARY; PRT; 923 AA.
ID Q90ZNS
AC Q90ZNS; 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 21, Last annotation update)
DE Eph receptor Eph3.
OS Brachydanio rerio (zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94200158; PubMed=8149909;
RA Xu Q., Holder N., Patient R., Wilson S.W.;
RT "Spatially regulated expression of three receptor tyrosine kinase
genes during gastrulation in the zebrafish."
RL Development 120:287-299(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Brennan C.H., Xu Q., Sordino P.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF369382; AAK54726.1;
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001090; Ephrin_receptor.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR001245; Tyr_pkinase.
DR InterPro: IPR001426; Ykrase_receptor.
DR Pfam; PF01404; EPH_lbd; 1.
DR Pfam; PF00041; fn3; 2.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR ProDom; PD001495; Ephrin_receptor; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
DR PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; UNKNOWN_1.
KW ATP-binding; Receptor; Transferase.
SQ SEQUENCE 923 AA; 101719 MW; 30DD8B1C87239790 CRC64;

Query Match 74.5%; Score 38; DB 13; Length 923;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
Db 360 CLHGGSC 367
:|||||

RESULT 5
Q9XWX5 PRELIMINARY; PRT; 1743 AA.
ID Q9XWX5
AC Q9XWX5; 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE Y43F8B.3 protein.
GN Y43F8B.3.
OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoida;
 OC Rhabditidae; Peloderinae; Caenorhabditis.

RN NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ainscough R.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL; AL032623; CAA21511.1; -;
 DR HSP; P05067; IAA
 DR InterPro; IPR002223; Kunitz_BPTI.
 DR InterPro; IPR002899; WRI/EB.
 DR Pfam; PF00014; Kunitz_BPTI; 15.
 DR PRINTS; PR00759; BASICPTASE.
 DR PRODOM; PD000222; Kunitz_BPTI; 15.
 DR SMART; SM00131; KU; 15.
 DR SMART; SM00289; WRI; 10.
 DR PROSITE; PS00280; BPTI_KUNITZ_1; 2.
 DR PROSITE; PS00279; BPTI_KUNITZ_2; 15.
 KW Serine protease inhibitor.
 SQ SEQUENCE 1743 AA; 191111 MW; B93C163556433C2A CRC64;

Query Match 74.58; Score 38; DB 5; Length 1743;
 Best Local Similarity 75.08; Pred. No. 58;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
 Db 1557 CARGNSC 1564

RESULT 6
 Q97H85
 ID Q97H85 PRELIMINARY; PRT; 452 AA.
 AC Q97H85;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE UDP-N-acetylmuramyl pentapeptide synthase.
 GN CAC2128.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 OX NCBI_TaxID=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Berton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 FT "Genome sequence and comparative analysis of the solvent-producing
 bacterium Clostridium acetobutylicum.";
 RL J. Bacteriol. 183:4823-4838(2001).
 DR EMBL; AE007713; AAK80086.1; -;
 DR InterPro; IPR00713; Mur_ligase.
 DR InterPro; IPR004101; Mur_ligase_C.
 DR Pfam; PF01225; Mur_ligase; 1.
 DR Pfam; PF02875; Mur_ligase_C; 1.
 DR TIGRFAMs; TIGR01143; murF; 1.
 KW Complete proteome.
 SQ SEQUENCE 452 AA; 50645 MW; A58ECD760732ABC5 CRC64;

Query Match 72.58; Score 37; DB 16; Length 452;
 Best Local Similarity 85.78; Pred. No. 27;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CLHRGNS 7
 Db 256 CLHEGNS 262

RESULT 7
 Q74832
 ID Q74832 PRELIMINARY; PRT; 524 AA.
 AC Q74832;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Putative mitochondrial translation system compon ent.
 GN SPC1183.04C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972H-;
 RA Hilbert H., Duesterhoeft A., Wood V., Rajandream M.A., Barrell B.G.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL031740; CAA21084.1; -;
 SQ SEQUENCE 524 AA; 60760 MW; FB4021F1B1C04290 CRC64;

Query Match 71.68; Score 36.5; DB 3; Length 524;
 Best Local Similarity 63.68; Pred. No. 39;
 Matches 7; Conservative 1; Mismatches 0; Indels 3; Gaps 1;

QY 1 CLHRG---NSC 8
 Db 20 CLHKGFGQINSC 30

RESULT 8
 P92546
 ID P92546 PRELIMINARY; PRT; 106 AA.
 AC P92546;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE ORF106G.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Unseld M., Marienfeld J.R., Brandt P., Brennick A.;
 RL Nat. Genet. 0:0-0(0).
 DR EMBL; Y08502; CAA69799.1; -;
 KW Mitochondrion.
 SQ SEQUENCE 106 AA; 12171 MW; 9972EFC5EFE206D4 CRC64;

Query Match 70.68; Score 36; DB 8; Length 106;
 Best Local Similarity 85.78; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LHRGNSC 8
 Db 2 LHRGRSC 8

RESULT 9
 P93299
 ID P93299 PRELIMINARY; PRT; 106 AA.
 AC P93299;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)

Wed Jul 16 13:41:38 2003

us-09-910-582b-9.rspt

DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE ORF106A (Hypothetical 12.2 kDa protein).

GN AT2G07701.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;
 [1]
 RN SEQUENCE FROM N.A.
 RP Unseqld M., Marlenfeld J.R., Brandt P., Brennicke A.;
 RL Nat. Genet. 0:0-0(0).
 [2]

RN SEQUENCE FROM N.A.
 RC STRAIN=CV. COLOMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M.-I., Greasy T.H., Haas B.J.,
 Wu D., Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
 Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RA "Arabidopsis thaliana chromosome 2 BAC T18C6 genomic sequence."
 RT Submitted (FEB-2002) to the EMBL/GenBank/DBSJ databases.

RL EMBL; Y08501; CAA69728.1; - - -
 DR EMBL; AC007729; AAM15500.1; - - -
 DR Hypothetical protein.
 KW SEQUENCE 106 AA; 12171 MW; 9972EFC5EFE206D4 CRC64;

Query Match 70.6%; Score 36; DB 10; Length 106;
 Best Local Similarity 85.7%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LHRGNSC 8
 :|||||
 DB 2 LHRGSC 8

RESULT 10
 Q8TTL2 PRELIMINARY; PRT; 179 AA.
 AC Q8TTL2;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Iron-sulfur flavoprotein.
 GN ISF-4 OR WA0418.
 OS Methanosarcina acetivorans.
 OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
 OC Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2214;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE=21929760; PubMed=11932238;
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
 Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
 Allen N., Naylor J., Stange-Thomann N., Dearrellano K., Johnson R.,
 RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Haderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 RA Perry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.;
 RA "The genome of Methanosarcina acetivorans reveals extensive metabolic
 RT and physiological diversity."
 RL Genome Res. 12:532-542(2002).
 DR EMBL; AE010702; AAM03866.1; - - -
 DR Complete proteome.
 KW SEQUENCE 179 AA; 19665 MW; EEF8CF8780537D41 CRC64;

Query Match 70.6%; Score 36; DB 17; Length 179;
 Best Local Similarity 62.5%; Pred. No. 19;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
 :|||||
 DB 55 CFRNGNC 62

RESULT 11

Q8TC85 PRELIMINARY; PRT; 327 AA.
 ID Q8TC85;
 AC Q8TC85;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Similar to RIKEN cDNA 4930453N24 gene.
 DE Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUP-TESTIS;
 RA Strausberg R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBSJ databases.
 RL EMBL; BC024188; AAH24188.1; - - - 539614E763F799B9 CRC64;
 SQ SEQUENCE 327 AA; 37280 MW;

Query Match 70.6%; Score 36; DB 4; Length 327;
 Best Local Similarity 71.4%; Pred. No. 32;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LHRGNSC 8
 :|||||
 DB 238 VHRGNTC 244

RESULT 12
 Q9NEC5 PRELIMINARY; PRT; 1396 AA.
 ID Q9NEC5;
 AC Q9NEC5;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Hypothetical 145.5 kDa protein.
 GN L8325.03.
 OS Leishmania major.
 OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=FRIEDLIN;
 RA Bothe G., Pohl T., Ivens A.C., Quail M., Rajandream M.A.,
 RA Bartell B.G.;
 RA Submitted (FEB-2000) to the EMBL/GenBank/DBSJ databases.
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=FRIEDLIN;
 RX MEDLINE=98146435; PubMed=9477341;
 RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
 RA Smith D.F.;
 RA "A physical map of the Leishmania major Friedlin genome."
 RT Genome Res. 8:135-145(1998).
 RL EMBL; AL157415; CAB75630.1; - - -
 DR Hypothetical protein.
 KW SEQUENCE 1396 AA; 145523 MW; E704D5F03909AA36 CRC64;

Query Match 70.6%; Score 36; DB 5; Length 1396;
 Best Local Similarity 62.5%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CLHRGNSC 8
 :|||||
 DB 1042 CRRHSGC 1049

RESULT 13

```

Q8WY30
ID Q8WY30 PRELIMINARY; PRT; 1731 AA.
AC
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Candidate tumor suppressor protein (Fragment).
GN LRP1B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20487554; PubMed=11031110;
RA Liu C.-X., Musco S., Lisitsyna N.M., Yaklichkin S.Y., Lisitsyn N.A.;
RT "Genomic organization of a new candidate tumor suppressor gene,
  LRP1B."
RL Genomics 69:271-274(2000).
DR EMBL; AF283374; AAL38108.1; JOINED.
DR EMBL; AF283344; AAL38108.1; JOINED.
DR EMBL; AF283345; AAL38108.1; JOINED.
DR EMBL; AF283346; AAL38108.1; JOINED.
DR EMBL; AF283347; AAL38108.1; JOINED.
DR EMBL; AF283348; AAL38108.1; JOINED.
DR EMBL; AF283349; AAL38108.1; JOINED.
DR EMBL; AF283350; AAL38108.1; JOINED.
DR EMBL; AF283351; AAL38108.1; JOINED.
DR EMBL; AF283352; AAL38108.1; JOINED.
DR EMBL; AF283353; AAL38108.1; JOINED.
DR EMBL; AF283354; AAL38108.1; JOINED.
DR EMBL; AF283355; AAL38108.1; JOINED.
DR EMBL; AF283356; AAL38108.1; JOINED.
DR EMBL; AF283357; AAL38108.1; JOINED.
DR EMBL; AF283358; AAL38108.1; JOINED.
DR EMBL; AF283359; AAL38108.1; JOINED.
DR EMBL; AF283360; AAL38108.1; JOINED.
DR EMBL; AF283361; AAL38108.1; JOINED.
DR EMBL; AF283362; AAL38108.1; JOINED.
DR EMBL; AF283363; AAL38108.1; JOINED.
DR EMBL; AF283364; AAL38108.1; JOINED.
DR EMBL; AF283365; AAL38108.1; JOINED.
DR EMBL; AF283367; AAL38108.1; JOINED.
DR EMBL; AF283368; AAL38108.1; JOINED.
DR EMBL; AF283369; AAL38108.1; JOINED.
DR EMBL; AF283370; AAL38108.1; JOINED.
DR EMBL; AF283371; AAL38108.1; JOINED.
DR EMBL; AF283372; AAL38108.1; JOINED.
DR EMBL; AF283373; AAL38108.1; JOINED.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000333; Ldl_receptor_rep.
DR InterPro; IPR002172; Ldl_receptor_A.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00057; ldl_recept_a; 9.
DR Pfam; PF00058; ldl_recept_b; 18.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00192; LDLa; 9.
DR SMART; SM00135; LY; 18.
DR PROSITE; PS01186; EGF_2; UNKNOWN_3.
DR PROSITE; PS01209; LDLRA_1; UNKNOWN_9.
DR PROSITE; PS00068; LDLRA_2; 9.
FT NON_TER
FT NON_TER
SQ SEQUENCE 1731 AA; 19384 MW; 56A007A83C490C07 CRC64;

Query Match 70.6%; Score 36; DB 4; Length 1731;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CLHRGNS 7
|||

```

Db 1234 CLYRGNS 1240

RESULT 14

```

Q9NZR2
ID Q9NZR2 PRELIMINARY; PRT; 4599 AA.
AC
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Low density lipoprotein receptor related protein-deleted in
  tumor.
DE LRPDIT.
GN LRPDIT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20227253; PubMed=10766186;
RA Liu C.-X., Musco S., Lisitsyna N.M., Forgacs E., Minna J.D.,
RA Lisitsyn N.A.;
RT "LRP-DIT, a putative endocytic receptor gene, is frequently
  inactivated in non-small cell lung cancer cell lines."
RL Cancer Res. 60:1961-1967(2000).
DR EMBL; AF176832; AAF70379.1; JOINED.
DR HSSP; Q07954; ICR8.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR InterPro; IPR002172; Ldl_recept_A.
DR InterPro; IPR001005; Myb_DNA_binding.
DR InterPro; IPR000169; Shprot_acsite.
DR Pfam; PF00008; EGF; 16.
DR Pfam; PF00057; ldl_recept_a; 32.
DR Pfam; PF00058; ldl_recept_b; 34.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00179; EGF_CA; 4.
DR SMART; SM00001; EGF_like; 18.
DR SMART; SM00192; LDLa; 32.
DR SMART; SM00135; LY; 33.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS00022; EGF_1; UNKNOWN_5.
DR PROSITE; PS01186; EGF_2; 9.
DR PROSITE; PS01187; EGF_CA; 3.
DR PROSITE; PS01209; LDLRA_1; 27.
DR PROSITE; PS00068; LDLRA_2; 32.
DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
DR PROSITE; PS00139; THIOLEPROTEASE_CYS; UNKNOWN_1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation;
KW Lipoprotein; Receptor; Repeat.
SQ SEQUENCE 4599 AA; 515496 MW; 9A97A555F1FAA866 CRC64;

Query Match 70.6%; Score 36; DB 4; Length 4599;
Best Local Similarity 85.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CLHRGNS 7
|||

```

RESULT 15

```

Q9JII8
ID Q9JII8 PRELIMINARY; PRT; 4599 AA.
AC
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Low density lipoprotein receptor related protein LRP1B/LRP-DIT.
GN LRP1B.
OS Mus musculus (Mouse).

```

wed Jul 16 13:41:38 2003

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Yaklichkin S., Lisitsyn N.;
RT "The structure of the mouse homologue of the human candidate tumor
RT suppressor gene LRP1B/LRP-DIT";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF270884; AAF81758.1;
DR HSSP; Q07954; 1CR8.
DR MGD; MGI:2151136; Lrp1b.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR000033; Ldl_receptor_rep.
DR InterPro: IPR002172; LDL_recept_A.
DR InterPro: IPR000169; SHprot_acsite.
DR Pfam; PF00008; EGF_17.
DR Pfam; PF00057; ldl_recept_a; 32.
DR Pfam; PF00058; ldl_recept_b; 33.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00179; EGF_CA; 4.
DR SMART; SM00001; EGF_like; 17.
DR SMART; SM00192; LDLa; 32.
DR SMART; SM00135; LY; 33.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS00022; EGF_1; UNKNOWN_5.
DR PROSITE; PS01186; EGF_2; 9.
DR PROSITE; PS01187; EGF_CA; 3.
DR PROSITE; PS01209; LDLRA_1; 27.
DR PROSITE; PS00068; LDLRA_2; 32.
DR PROSITE; PS00139; THIOLEPROTEASE_CYS; UNKNOWN_1.
DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation;
KW Lipoprotein; Receptor; Repeat.
SQ SEQUENCE 4599 AA; 513623 MW; 11462A3354FFB200 CRC64;

Query Match 70.68; Score 36; DB 11; Length 4599;
Best Local Similarity 85.7%; Pred. No 3.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CLHRGNS 7
DB 2158 CLYRGNS 2164
|||||

Search completed: July 16, 2003, 13:15:14
Job time : 22.6585 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:04:13 ; Search time 40.9756 Seconds
(without alignments)
39.023 Million cell updates/sec

Title: US-09-910-582b-10
Perfect score: 74
Sequence: 1 CRSWNKADNRSC 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	100.0	12	22	Peptide which sele
2	74	100.0	12	22	Heart homing pepi
3	46	62.2	101	23	Human ORF1511 prot
4	42	56.8	82	22	Propionibacterium
5	42	56.8	112	15	Light chain variab
6	42	56.8	575	22	Novel human diagno
7	42	56.8	575	22	Novel human diagno
8	41	55.4	170	21	Pinus radiata cell
9	41	55.4	998	21	Pinus radiata cell
10	40.5	54.7	439	23	Human DITHP polype

11	54.1	51	22	AAU48787	Propionibacterium
12	40	84	22	ABG10762	Novel human diagno
13	40	1045	22	ABB62186	Drosophila melanog
14	39	52.7	49	22	Plant defensin pep
15	39	52.7	50	16	Aesculus hippocast
16	39	52.7	50	18	Aesculus hippocast
17	39	52.7	50	20	Antifungal peptide
18	39	52.7	136	21	zea mays protein f
19	39	52.7	144	21	pinus radiata cell
20	39	52.7	228	21	Arabidopsis thalia
21	39	52.7	230	21	Arabidopsis thalia
22	39	52.7	286	23	Lactococcus lactis
23	39	52.7	318	21	Lactococcus lactis
24	39	52.7	409	23	Herbicidally activ
25	39	52.7	482	21	Arabidopsis thalia
26	39	52.7	505	21	Arabidopsis thalia
27	39	52.7	518	22	Novel human diagno
28	39	52.7	547	21	Pinus radiata cell
29	39	52.7	1337	22	Human protein SEQ
30	39	52.7	1338	22	Human protein SEQ
31	38	51.4	15	23	C textile mu-conop
32	38	51.4	16	23	C gloriamaris mu-c
33	38	51.4	18	15	MtRiib-irra-33, bi
34	38	51.4	49	15	Antimicrobial Ct-A
35	38	51.4	49	20	Antifungal peptide
36	38	51.4	49	23	Transplant media a
37	38	51.4	50	14	Antimicrobial prot
38	38	51.4	64	23	Human ORFX protein
39	38	51.4	66	22	Human immune/haema
40	38	51.4	70	23	C textile mu-conop
41	38	51.4	71	23	C gloriamaris mu-c
42	38	51.4	92	22	Propionibacterium
43	38	51.4	94	22	Novel human diagno
44	38	51.4	235	22	Lactococcus lactis
45	38	51.4	245	22	Human peptide #282

ALIGNMENTS

RESULT 1
AAB30899 standard; peptide; 12 AA.
ID AAB30899 standard; peptide; 12 AA.
XX
AC AAB30899;
XX
DT 02-APR-2001 (first entry)
XX
DE Peptide which selectively binds to normal cardiac endothelium.
XX
KW Cardiac endothelium; angiogenic factor; vascular endothelium;
KW peripheral vascular disease; cardiovascular disease; angiogenesis;
KW cardiac neovascularisation.
XX
OS Unidentified.
XX
PN WO200075329-A1.
XX
PD 14-DEC-2000.
XX
PF 31-MAY-2000; 2000WO-US14988.
XX
PR 07-JUN-1999; 99US-0327045.
XX
PA (EDWA-) EDWARDS LIFESCIENCES CORP.
XX (BAXT) BAXTER AG.
PI Levine AJ, Mitterer A, Falkner F, Scheifflinger F, Dörner F;
XX WPI; 2001-091212/10.
XX
PT New chimeric molecules having an angiogenic factor linked to a
PT targeting molecule that binds to a vascular endothelium, useful for

PT increasing cardiac neovascularisation, or treating peripheral vascular
XX and cardiovascular diseases
PS Disclosure: Page 27; 67pp; English.
XX AAB30895-99 represent targeting molecules, which are used to produce the
CC chimeric molecules of the invention. AAB30895-98 selectively bind to
CC normal cardiac endothelium. The specification describes a chimeric
CC molecule comprising an angiogenic factor linked to a targeting molecule
CC that specifically binds to a vascular endothelium. The chimeric
CC molecules are useful for treatment of peripheral vascular or
CC cardiovascular diseases. Specifically, they are useful for inducing or
CC inhibiting angiogenesis, for increasing cardiac neovascularisation in
CC ischemic tissue in the peripheral vascular system.
XX
XX SQ Sequence 12 AA;
Query Match 100.0%; Score 74; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Caps 0;
QY 1 CRSWNKADNRSC 12
Db 1 CRSWNKADNRSC 12
RESULT 2
ID AAB59304 standard; peptide; 12 AA.
XX AAB59304;
AC AAB59304;
XX 21-MAR-2001 (first entry)
DT Heart homing peptide SEQ ID NO: 10.
DE Heart homing peptide; cardiovascular disease; ischaemic disease;
XX gene therapy.
KW Synthetic.
OS Key Location/Qualifiers
FH Disulfide-bond 1..12
FT /note= "disulfide bond cyclises the peptide"
FT
XX WO200075174-A1.
XX 14-DEC-2000.
XX 31-MAY-2000; 2000WO-US15088.
XX 07-JUN-1999; 99US-0326718.
XX (BURN-) BURNHAM INST.
XX Ruoslahti E, Mackenna DA;
XX WPI; 2001-071059/08.
XX Novel heart homing peptide that selectively homes to normal ischaemic
XX and cardiac tissue useful for targeting ischaemic tissues for treating
PT ischaemic and cardiovascular diseases such as atherosclerosis and
PT restenosis
PT
XX Claim 2; Page 55; 70pp; English.
XX The present invention provides a number of heart homing peptides which
XX selectively home to cardiac tissue. These can be used in the treatment of
CC cardiovascular and ischaemic diseases, such as atherosclerosis,
CC thrombosis, restenosis, vasculitis, atherosclerotic aneurysms, myocardial
CC hypertrophy, congenital heart diseases, ischaemic heart disease and
CC anginas, acquired valvular/endocardial diseases, primary myocardial
CC diseases, cardiac tumours and arrhythmias.
CC

XX SQ Sequence 12 AA;
Query Match 100.0%; Score 74; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CRSWNKADNRSC 12
Db 1 CRSWNKADNRSC 12
RESULT 3
ID ABP32538 standard; protein; 101 AA.
XX ABP32538;
AC ABP32538;
XX 08-JUL-2002 (first entry)
DT Human ORF1511 protein, SEQ ID NO:3022.
DE Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
XX disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibitor; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
KW vasotrophic; antipsoriatic; antidiabetic; cytostatic; nootropic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.
XX Homo sapiens.
XX WO200190366-A2.
XX 29-NOV-2001.
XX 24-MAY-2001; 2001WO-US17076.
XX 24-MAY-2000; 2000US-206690P.
XX (CURA-) CURAGEN CORP.
XX Leach MD, Shimkets RA;
XX WPI; 2002-106200/14.
XX N-PSDB; ABN76564.
XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and disorders related to organ
PT transplantation
PT
XX Claim 10; Page 1014-1015; 2508pp; English.
XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
CC ABN79587 represent cDNAs encoding them. The invention also encompasses
CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC polynucleotides, the recombinant production of ORFX proteins, antibodies
CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
CC polypeptides, methods of screening for modulators of ORFX expression or
CC activity, and methods of screening individuals for a predisposition to an
CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC range of biological activities, such as cytokine, cell proliferation,
CC cell differentiation, immune modulation, haematopoiesis regulation,
CC

CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 CC and antiinfective activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.
 XX
 SQ Sequence 101 AA;

Query Match 62.2%; Score 46; DB 23; Length 101;
 Best Local Similarity 63.6%; Pred. No. 6;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CRSNKADNRS 11
 Db 23 CEANRNRNS 33
 I:|||||

RESULT 4
 AAU65366
 ID AAU65366 standard; Protein; 82 AA.

XX AC AAU65366;

XX DT 27-FEB-2002 (first entry)

XX DE Propionibacterium acnes immunogenic protein #26262.

XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

XX OS Propionibacterium acnes.

XX PN WO200181581-A2.

XX PD 01-NOV-2001.

XX PF 20-APR-2001; 2001WO-US12865.

XX PR 21-APR-2000; 2000US-199047P.

XX PR 02-JUN-2000; 2000US-208841P.

XX PR 07-JUL-2000; 2000US-216747P.

XX PA (CORI-) CORIXA CORP.

XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX DR WPI; 2001-616774/71.

XX DR N-PSDB; AAS99667.

XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -

XX PS Example 1; SEQ ID No 26561; 1069pp; English.

XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 82 AA;

Query Match 56.8%; Score 42; DB 22; Length 82;
 Best Local Similarity 58.3%; Pred. No. 21;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CRSNKADNRS 12
 Db 10 CRTRKPADNRC 21
 ||: |||||

RESULT 5
 AAR52035
 ID AAR52035 standard; Protein; 112 AA.

XX AC AAR52035;

XX DT 26-SEP-1996 (first entry)

XX DE Light chain variable region of murine antibody 2fb4.

XX KW antibody; humanised; murine; human; heavy chain; light; variable;
 KW framework region; complementarity determining region; reshaping;
 KW modelling; surface residue; modify.

XX OS Mus sp.

XX FH Key Location/Qualifiers

XX FT Region 1..22 /label= framework_region_1

XX FT /note= "FR 1"

XX FT Region 23..35 /label= complementarity_determining_region_1

XX FT /note= "CDR 1"

XX FT Region 36..50 /label= FR_2

XX FT Region 51..57 /label= CDR_2

XX FT Region 58..89 /label= FR_3

XX FT Region 90..99 /label= CDR_3

XX FT Region 100..112 /label= FR_4

XX PN EP592106-A1.

XX PD 13-APR-1994.

XX PF 07-SEP-1993; 93EP-0307051.

us-09-910-582b-10.rag

Wed Jul 16 13:41:33 2003

PR 09-SEP-1992; 92US-0942245.
 XX (PEDE/) PEDERSEN J T.
 PA (IMMU-) IMMUNOGEN INC.
 XX Guild BC, Pedersen JT, Rees AR, Roguska MA, Searle SMJ;
 XX WPI; 1994-120230/15.
 DR Method of resurfacing of rodent antibodies to produce humanised
 XX antibody forms - for producing non-human antibodies with improved
 PT therapeutic efficiency by presenting human surface on V-region
 XX Example 1; Fig 3A; 230pp; English.
 XX The present sequence is that of the light chain variable (LC VR) region
 CC of murine antibody 2fB4. This sequence was aligned with 11 other known
 CC antibody LC VRs and a set of framework positions of surface exposed amino
 CC acid residues was determined. This information can be used in a method to
 CC determine how to modify a rodent antibody or fragment by resurfacing in
 CC order to produce a humanised rodent antibody. Residues determined from
 CC alignment at positions 1, 3, 5, 9, 15, 18, 46, 47, 51, 63, 66, 73, 86,
 CC 87, 111, 115, 116 and 117 are accessible residues. None of the entire
 CC combinations of surface residues in the murine sequences were found in
 CC the human sequences and vice versa. However the residues in individual
 CC positions appear to be conserved.
 XX Sequence 112 AA;
 SQ Query Match 56.8%; Score 42; DB 15; Length 112;
 Best Local Similarity 66.7%; Pred. No. 29;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CRSWNKADN 9
 DB 89 CASWNSSDN 97
 I I I I I I I
 RESULT 6
 ABG17859
 ID ABG17859 standard; Protein: 575 AA.
 XX AC ABG17859;
 XX 18-FEB-2002 (first entry)
 DT Novel human diagnostic protein #17850.
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 KW Homo sapiens.
 XX WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 XX 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 XX N-PSDB; AAS82046.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

XX Claim 20; SEQ ID No 48218; 103pp; English.
 PS The invention relates to isolated polynucleotide (I) and
 XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG03077 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 575 AA;
 SQ Query Match 56.8%; Score 42; DB 22; Length 575;
 Best Local Similarity 50.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CRSWNKADNRC 12
 DB 144 CRSWSRSNYWSC 155
 I I I I I I I I
 RESULT 7
 ABG17888
 ID ABG17888 standard; Protein: 575 AA.
 XX AC ABG17888;
 XX 18-FEB-2002 (first entry)
 DT Novel human diagnostic protein #17879.
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 KW Homo sapiens.
 XX WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 XX 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 XX N-PSDB; AAS82075.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX Claim 20; SEQ ID No 48247; 103pp; English.
 PS

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 575 AA;

Query Match 56.8%; Score 42; DB 22; Length 575;
 Best Local Similarity 50.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 CRSWNKADNRSC 12

DB 144 CRSWSRSNYWSC 155

RESULT 8

AAB25450
 ID AAB25450 standard; Protein; 170 AA.

XX AAB25450;

XX 27-NOV-2000 (first entry).

DE Pinus radiata cell signalling involved protein SEQ ID NO:769.

XX Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
 KW plant cell signalling; modulation; transgenic plant; pathogen; growth;
 KW environmental change; development; cell proliferation; differentiation;
 KW elongation; survival; disease resistance; nutrient metabolism.

OS Pinus radiata.

XX WO2000042171-A1.

XX 20-JUL-2000.

XX 11-JAN-2000; 2000WO-US00724.

XX 12-JAN-1999; 99US-0228986.

PR 01-NOV-1999; 99US-0162866.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX Strabala TJ, Nieuwenhuizen NJ;

XX WPI; 2000-476052/41.

XX Isolated polynucleotide encoding a polypeptide involved in cell
 PT signaling used for generating transgenic plants with modified responses
 to external signals -

PS Claim 3; Page 356-357; 527pp; English.

XX AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide

CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or
 CC pine (Pinus radiata also known as Monterey pine). The protein sequences
 CC are involved in cell signalling. The polynucleotide and protein
 CC sequences can be used to modify the response of plant cells to external
 CC signals e.g. environmental changes or pathogens during the growth and
 CC development of a plant. They can be used to modify cell proliferation,
 CC differentiation, elongation and survival, resistance to disease and
 CC nutrient metabolism. Examples of modifications which can be produced are
 CC altered fruit ripening and senescence of leaves and flowers e.g. to
 CC delay senescence and prolong the life of cut flowers or enhance
 CC senescence of reproductive organs to engineer sterile plants. Other
 CC modifications can be used to delay senescence in selected cell types or
 CC organs providing fruit and vegetables which have a longer shelf life
 CC between harvest and consumption, or to decrease branching frequency in
 CC forest tree species giving long stretches of valuable knot-free clear
 CC wood which can be used in solid timber furniture and veneers.

XX Sequence 170 AA;

Query Match 55.4%; Score 41; DB 21; Length 170;

Best Local Similarity 54.5%; Pred. No. 64;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 RSWNKADNRSC 12

DB 48 RWNENDNSPC 58

RESULT 9

AAB25556

ID AAB25556 standard; Protein; 998 AA.

XX AAB25556;

XX 27-NOV-2000 (first entry)

DE Pinus radiata cell signalling involved protein SEQ ID NO:931.

XX Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
 KW plant cell signalling; modulation; transgenic plant; pathogen; growth;
 KW environmental change; development; cell proliferation; differentiation;
 KW elongation; survival; disease resistance; nutrient metabolism.

OS Pinus radiata.

XX WO2000042171-A1.

XX 20-JUL-2000.

XX 11-JAN-2000; 2000WO-US00724.

XX 12-JAN-1999; 99US-0228986.

PR 01-NOV-1999; 99US-0162866.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX Strabala TJ, Nieuwenhuizen NJ;

XX WPI; 2000-476052/41.

XX Isolated polynucleotide encoding a polypeptide involved in cell
 PT signaling used for generating transgenic plants with modified responses
 to external signals -

PS Claim 3; Page 505-507; 527pp; English.

XX AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide
 CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or
 CC pine (Pinus radiata also known as Monterey pine). The protein sequences
 CC are involved in cell signalling. The polynucleotide and protein
 CC sequences can be used to modify the response of plant cells to external
 CC signals e.g. environmental changes or pathogens during the growth and
 CC development of a plant. They can be used to modify cell proliferation,

Wed Jul 16 13:41:33 2003

CC differentiation, elongation and survival, resistance to disease and
CC nutrient metabolism. Examples of modifications which can be produced are
CC altered fruit ripening and senescence of leaves and flowers e.g. to
CC delay senescence and prolong the life of cut flowers or enhance
CC senescence of reproductive organs to engineer sterile plants. Other
CC modifications can be used to delay senescence in selected cell types or
CC organs providing fruit and vegetables which have a longer shelf life
CC between harvest and consumption, or to decrease branching frequency in
CC forest tree species giving long stretches of valuable knot-free clear
CC wood which can be used in solid timber furniture and veneers.

XX Sequence 998 AA;
SQ Query Match 55.4%; Score 41; DB 21; Length 998;
Best Local Similarity 54.5%; Pred. No. 4.1e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RSNWKNDRSC 12
Db 48 RSNWENDNSPC 58

RESULT 10
ABG60008
ID ABG60008 standard; Protein: 439 AA.

XX AC ABG60008;

XX 30-JUL-2002 (first entry)

XX Human DITHP polypeptide #66.

XX Human; DITHP; diagnostic and therapeutic polypeptide; bone; testis; skin;
KW cell proliferative disorder; cancer; tumour; autoimmune disorder; brain;
KW inflammatory disorder; viral infection; bacterial infection; seizure;
KW fungal infection; parasitic infections; developmental disorder; breast;
KW endocrine disorder; metabolic disorder; neurological disorder; cervix;
KW gastrointestinal disorder; transport disorder; gene therapy; kidney;
KW adrenal gland; bone marrow; lung; ovary; pancreas; prostate; spleen;
KW thymus.

XX Homo sapiens.

XX WO200220754-A2.

XX 14-MAR-2002.

XX 29-AUG-2001; 2001WO-US27127.

XX 05-SEP-2000; 2000US-229747P.

XX 05-SEP-2000; 2000US-229748P.

XX 05-SEP-2000; 2000US-229749P.

XX 05-SEP-2000; 2000US-229750P.

XX 05-SEP-2000; 2000US-229751P.

XX 05-SEP-2000; 2000US-230583P.

XX 06-SEP-2000; 2000US-230585P.

XX 06-SEP-2000; 2000US-230586P.

XX 06-SEP-2000; 2000US-230587P.

XX 06-SEP-2000; 2000US-230588P.

XX 07-SEP-2000; 2000US-230589P.

XX 07-SEP-2000; 2000US-230590P.

XX 07-SEP-2000; 2000US-231167P.

XX (INCY-) INCYTE GENOMICS INC.

XX PA

XX

PI Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL;
PI Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Roseberry AM;
PI Momiya MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;
PI Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V, Daffo A;
PI Marwaha R, Chen AJ, Chang SC, Au AP, Inman RR;

DR WPI: 2002-383054/41.

DR N-PSDB: ABK71599.

XX An isolated polynucleotide useful in diagnostics and therapeutics -

XX Claim 29; Page 569-570; 686pp; English.

XX The invention relates to human diagnostic and therapeutic (dithp)
CC polynucleotides and their associated polypeptides (DITHP polypeptides).
CC The sequences of the invention are used in the treatment and diagnosis of
CC cell proliferative disorders (e.g. atherosclerosis, cirrhosis), cancers
CC (e.g. tumours of the adrenal gland, bone, bone marrow, brain, breast,
CC cervix, kidney, lung, ovary, pancreas, prostate, skin, spleen, testis or
CC thymus), autoimmune/inflammatory disorders (e.g. asthma, bronchitis,
CC psoriasis, osteoporosis), viral infections, bacterial infections, fungal
CC infections, parasitic infections, developmental disorders (e.g. anaemia,
CC epilepsy), seizure disorders (e.g. thrombosis, aneurysm), metabolic disorders
CC endocrine disorders (e.g. diabetes), neurological disorders (e.g. stroke,
CC (e.g. obesity, diabetes), multiple sclerosis), gastrointestinal
CC amyotrophic lateral sclerosis, colitis, lysinuria) and transport disorders
CC disorders (e.g. ulcerative colitis, catatonias, peripheral neuropathy). Sequences
CC (e.g. myotonic dystrophy, catatonias, peripheral neuropathy). Sequences
CC ABG59943-ABG60220 represent human DITHP polypeptides of the invention.

XX Sequence 439 AA;

XX Query Match 54.7%; Score 40.5; DB 23; Length 439;

XX Best Local Similarity 58.3%; Pred. No. 2.1e+02; Indels 1; Gaps 1;

XX Matches 7; Conservative 2; Mismatches 2;

QY 1 CRSWKNDRSC 12

Db 22 CRTWSSIDN-SC 32

RESULT 11

AAU48787

ID AAU48787 standard; Protein: 51 AA.

XX AC AAU48787;

XX 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #9683.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

XX 02-JUN-2000; 2000US-208841P.

XX 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.
 DR N-PSDB; AAS59543.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris .
 XX
 XX Example 1; SEQ ID NO 9982; 1069pp; English.
 PS
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system; however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 51 AA;

Query Match 54.1%; Score 40; DB 22; Length 51;
 Best Local Similarity 54.5%; Pred. No. 26;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CRSNKADNRS 11
 |||| : : ||
 DB 29 CRSGRHEQRS 39

RESULT 12
 ABG10762
 ID ABG10762 standard; Protein; 84 AA.

XX AC ABG10762;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #10753.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS74949.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity .
 XX

PS Claim 20; SEQ ID NO 41121; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 84 AA;

Query Match 54.1%; Score 40; DB 22; Length 84;
 Best Local Similarity 54.5%; Pred. No. 44;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RSNKADNRSC 12

DB 55 RAWQQLDNGSC 65

RESULT 13
 ABB62186
 ID ABB62186 standard; Protein; 1045 AA.

XX AC ABB62186;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 13350.

XX KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL06289.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell

us-09-910-582b-10.rag

Wed Jul 16 13:41:33 2003

PT interactions -
PS Disclosure; SEQ ID NO 13350; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (AB857737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 1045 AA;
Query Match 54.1%; Score 40; DB 22; Length 1045;
Best Local Similarity 60.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 3 SWNKADNRSC 12
DB 792 SWGSGDNRCC 801
RESULT 14
AAB04047
ID AAB04047 standard; Protein; 49 AA.
XX
XX AAB04047;
DT 11-APR-2001 (first entry)
XX
DE Plant defensin peptide fragment.
XX
XX Defensin; disease; fungus; resistance; transgenic plant; vaccine;
KW immunisation; antibody; crop protection.
XX
XX Picramnia pentandra.
XX
XX Key Location/Qualifiers
FH Misc-difference 8 /note= "Ile, Thr, Lys or Arg"
FT Misc-difference 22 /note= "Gln or His"
FT Misc-difference 32 /note= "Arg or Gly"
FT Misc-difference 46 /note= "Phe, Leu, Ile or Val"
FT
XX WO200068405-A2.
PN
XX 16-NOV-2000.
XX
XX 03-MAY-2000; 2000WO-US11952.
XX
XX 07-MAY-1999; 99US-0133039.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX
XX Miao G, Weng Z, Famodu OO;
XX WPI; 2001-024871/03.
XX N-PSDB; AAA54388.
XX
XX Isolated polynucleotides encoding plant defensins, useful for creating
XX transgenic plants with higher or lower levels of defensin polypeptides,
XX especially for increasing disease (e.g. fungal) resistance and stress
XX tolerance
XX
XX Claim 2; Page 39; 50pp; English.
PS
XX Nucleotides encoding plant defensins can be used to create
CC transgenic plants in which plant defensin polypeptides are present
CC at higher or lower levels than normal, or in cell types or
CC developmental stages in which they are not normally found. This has
CC the effect of altering the level of disease (e.g. fungal) resistance
CC and stress tolerance in those cells. The defensin polypeptides are
CC useful for immunising animals to produce polyclonal or monoclonal
CC antibodies. These antibodies are useful for screening cDNA
CC expression libraries to isolate full-length plant defensin cDNA
CC clones.
XX
XX SQ Sequence 49 AA;
Query Match 52.7%; Score 39; DB 22; Length 49;
Best Local Similarity 41.7%; Pred. No. 36;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 CRSWNKADNRSC 12
DB 23 CKSWEHAQHAC 34
RESULT 15
AAR76530
ID AAR76530 standard; protein; 50 AA.
XX
XX AAR76530;
AC
XX 15-FEB-1996 (first entry)
XX
DE Aesculus hippocastanum antimicrobial protein 1.
XX
XX Antifungal; antimicrobial; antibiotic; Ah-AMP1; horse chestnut;
KW fungal resistant plant; crop protection; food preservative.
XX
XX Aesculus hippocastanum.
OS
XX
XX Key Location/Qualifiers
FH Protein 1..50
FT /label= Ah-AMP1
FT /note= "includes 8 conserved cysteine residues at
FT positions 2, 14, 20, 24, 35, 44, 46 and 50,
FT two conserved glycines at positions 12 and
FT 33, a conserved aromatic residue (trp) at
FT position 10 and a conserved Glu at position
FT 28"
XX
XX WO9518229-A1.
XX
XX 06-JUL-1995.
PD
XX
XX 19-DEC-1994; 94WO-GB02766.
PF
XX
XX 24-DEC-1993; 93GB-0026424.
XX
XX (ZENE) ZENECA LTD.
XX
XX Broekaert WF, Cammue BPA, Osborn RW, Rees SB;
XX WPI; 1995-246394/32.
XX
XX New antimicrobial proteins from Heuchera and Aesculus seeds - for
XX treating infections in humans, in plant protection and as food
XX preservative also related DNA and transformed microorganisms and
XX plants
XX
XX Claim 1; Page 29 and Fig 5; 45pp; English.
XX
XX An antimicrobial protein (Ah-AMP1) was extracted from the seeds of
XX Aesculus hippocastanum. The amino acid sequence of Ah-AMP1 (AAR76530)
XX is homologous to the known Rs-AFP-type antifungal/antimicrobial
XX proteins, sharing the consensus cysteine-glycine motif and having a

CC total of 8 conserved cysteine residues. The antifungal activity of
CC Ah-AMPI does not cause hyperbranching of fungal hyphae. The protein
CC also has antibiotic activity and can be used, e.g. as a
CC preservative in foods.

XX

SQ Sequence 50 AA;

Query Match 52.7%; Score 39; DB 16; Length 50;

Best Local Similarity 41.7%; Pred. No. 37;

Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CRSWNKADNRSC 12

I : I I I : I

Db 24 CODWEKASHGAC 35

Search completed: July 16, 2003, 13:12:42

Job time : 42.9756 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:10:13 ; Search time 14.0488 seconds
(without alignments)
25.132 Million cell updates/sec

Title: US-09-910-582b-10
Perfect score: 74
Sequence: 1 CRSWNKADNRSC 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2.6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2.6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2.6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2.6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2.6/prodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2.6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	100.0	12	4	US-09-326-718-10
2	42	56.8	112	1	US-07-942-245-6
3	41	55.4	107	6	5242821-7
4	39	52.7	50	1	US-08-656-318A-2
5	39	52.7	50	2	US-08-956-459-2
6	38	51.4	18	5	PCT-US94-01234-33
7	38	51.4	49	1	US-08-377-687-15
8	38	51.4	49	1	US-08-377-687-24
9	38	51.4	49	1	US-08-656-318A-7
10	38	51.4	49	2	US-08-777-192-15
11	38	51.4	49	2	US-08-777-192-24
12	38	51.4	49	2	US-08-956-459-7
13	38	51.4	49	4	US-08-971-982-15
14	38	51.4	49	4	US-08-971-982-24
15	38	51.4	558	4	US-09-199-637A-277
16	38	51.4	978	2	US-08-415-593-43
17	37	50.0	18	5	PCT-US94-01234-25
18	37	50.0	192	4	US-09-475-316A-25
19	37	50.0	192	4	US-09-475-316A-87
20	37	50.0	493	4	US-09-112-498A-2
21	37	50.0	1182	4	US-09-287-354-6
22	36	48.6	1276	4	US-08-937-236-3
23	36	48.6	1277	4	US-08-937-236-6
24	36	48.6	1291	4	US-08-569-214-3
25	36	48.6	1291	4	US-08-937-236-2
26	36	48.6	1292	4	US-08-569-214-5
27	36	48.6	1292	4	US-08-569-214-6

28 48.6 1292 4 US-08-937-236-5 Sequence 5, Appli
29 48.6 1295 4 US-08-569-214-2 Sequence 2, Appli
30 47.3 33 4 US-09-172-841-29 Sequence 29, Appli
31 47.3 50 1 US-08-377-687-10 Sequence 10, Appli
32 47.3 50 1 US-08-377-687-12 Sequence 12, Appli
33 47.3 50 1 US-08-377-687-13 Sequence 13, Appli
34 47.3 50 1 US-08-377-687-20 Sequence 20, Appli
35 47.3 50 1 US-08-377-687-21 Sequence 21, Appli
36 47.3 50 1 US-08-377-687-22 Sequence 22, Appli
37 47.3 50 1 US-08-656-318A-5 Sequence 5, Appli
38 47.3 50 1 US-08-656-318A-6 Sequence 6, Appli
39 47.3 50 2 US-08-777-192-10 Sequence 10, Appli
40 47.3 50 2 US-08-777-192-12 Sequence 12, Appli
41 47.3 50 2 US-08-777-192-13 Sequence 13, Appli
42 47.3 50 2 US-08-777-192-20 Sequence 20, Appli
43 47.3 50 2 US-08-777-192-21 Sequence 21, Appli
44 47.3 50 2 US-08-777-192-22 Sequence 22, Appli
45 47.3 50 2 US-08-956-459-5 Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-326-718-10
; Sequence 10, Application US/09326718
; Patent No. 6303573
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; TITLE OF INVENTION: Heart Homing Peptides and Methods of
; TITLE OF INVENTION: Using Same
; FILE REFERENCE: P-LJ 3512
; CURRENT APPLICATION NUMBER: US/09/326,718
; CURRENT FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-09-326-718-10

Query Match 100.0%; Score 74; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRSWNKADNRSC 12

Db 1 CRSWNKADNRSC 12

RESULT 2
US-07-942-245-6
; Sequence 6, Application US/07942245
; Patent No. 5639641
; GENERAL INFORMATION:
; APPLICANT: PEDERSEN, Jan T.
; APPLICANT: SEARLE, Stephen M.J.
; APPLICANT: REES, Anthony R.
; APPLICANT: ROGUSKA, Michael A.
; APPLICANT: GUILD, Braydon C.
; TITLE OF INVENTION: SURFACE RESIDUE VENERING OF RODENT
; TITLE OF INVENTION: ANTIBODIES
; NUMBER OF SEQUENCES: 522
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States

us-09-910-582b-10.ra1

Wed Jul 16 13:41:33 2003

ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: HP 9000/700 Workstation
OPERATING SYSTEM: UNIX
SOFTWARE: In house
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/942,245
FILING DATE: 09-SEP-1992
CLASSIFICATION: 530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 112 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-942-245-6

Query Match 56.8%; Score 42; DB 1; Length 112;
Best Local Similarity 66.7%; Pred. No. 10;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CRSNKADN 9
| | | | |
Db 89 CASWSSDN 97

RESULT 3
5242821-7
Patent No. 5242821
APPLICANT: PALVA, LIKKA-SIBAKOV, MERVIL
TITLE OF INVENTION: LACTOCOCCUS PROMOTER AND SIGNAL
SEQUENCES FOR EXPRESSION IN BACTERIA
NUMBER OF SEQUENCES: 27
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/377,450
FILING DATE: 10-JUL-1989
SEQ ID NO: 7:
LENGTH: 107
5242821-7

Query Match 55.4%; Score 41; DB 6; Length 107;
Best Local Similarity 70.0%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RSNKADNRS 11
| | | | |
Db 86 RSWNDSSNR 95

RESULT 4
US-08-656-318A-2
Sequence 2, Application US/08656318A
Patent No. 5750504
GENERAL INFORMATION:
APPLICANT: BROEKAERT, WILLEM F.
APPLICANT: CAMMUE, BRUNO P.A.
APPLICANT: OSBORN, RUPERT W.
APPLICANT: REES, SARAH B.
TITLE OF INVENTION: ANTIMICROBIAL PROTEINS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DABY & CUSHMAN
ADDRESSEE: Intellectual Property Group of
ADDRESSEE: PILLSBURY MADISON & SUTRO LLP
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.

ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/656,318A
FILING DATE: 12-JUN-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9326424.0
FILING DATE: 24-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB94/02766
FILING DATE: 19-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 224199/SEE37925/UST
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3075
TELEFAX: (202) 822-0944
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: AH-AMP1
US-08-656-318A-2

Query Match 52.7%; Score 39; DB 1; Length 50;
Best Local Similarity 41.7%; Pred. No. 13;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CRSNWKADNRSC 12
| | | | |
Db 24 CQWKEKASHGAC 35

RESULT 5
US-08-936-459-2
Sequence 2, Application US/08956459
Patent No. 5919918
GENERAL INFORMATION:
APPLICANT: BROEKAERT, WILLEM F.
APPLICANT: CAMMUE, BRUNO P.A.
APPLICANT: OSBORN, RUPERT W.
APPLICANT: REES, SARAH B.
TITLE OF INVENTION: ANTIMICROBIAL PROTEINS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY MADISON & SUTRO LLP
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,459
FILING DATE: 22-OCT-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/656,318
FILING DATE: 12-JUN-1996

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB94/02766
; FILING DATE: 19-DEC-1994
; PRIOR APPLICATION DATA: GB 9326424.0
; FILING DATE: 24-DEC-1993
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Ah-AMP1
; US-08-956-459-2

Query Match 52.7%; Score 39; DB 2; Length 50;
Best Local Similarity 41.7%; Pred. No. 13;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CRSNKADNRSC 12
I: ||||:|
Db 24 CQWKRASHGAC 35

RESULT 6
PCT-US94-01234-33
; Sequence 33, Application PC/TUS9401234
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE
; TITLE OF INVENTION: BINDING SITES
; NUMBER OF SEQUENCES: 76
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01234
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/084,542
; FILING DATE: 28-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/012,566
; FILING DATE: 02-FEB-1993
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; PCT-US94-01234-33

Query Match 51.4%; Score 38; DB 5; Length 18;
Best Local Similarity 50.0%; Pred. No. 7;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 SMNKADNRSC 12
I: :|||:|
Db 4 SFGADNRNC 13

RESULT 7
US-08-377-687-15
; Sequence 15, Application US/08377687
; Patent No. 5538525
; GENERAL INFORMATION:
; APPLICANT: BROEKERT, WILLEM F.
; APPLICANT: CAMMUE, BRUNO P.A.

; APPLICANT: REES, SARAH B.
; APPLICANT: TERRAS, FRANKY R.G.
; APPLICANT: VANDERLEYDEN, JOZEF
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/377,687
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/002,480
; FILING DATE: 04-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-377-687-15

Query Match 51.4%; Score 38; DB 1; Length 49;
Best Local Similarity 41.7%; Pred. No. 18;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CRSNKADNRSC 12
I: ||||:|
Db 24 CRWESAKHGAC 35

RESULT 8
US-08-377-687-24
; Sequence 24, Application US/08377687
; Patent No. 5538525
; GENERAL INFORMATION:
; APPLICANT: BROEKERT, WILLEM F.
; APPLICANT: CAMMUE, BRUNO P.A.
; APPLICANT: OSBORN, RUPERT W.
; APPLICANT: REES, SARAH B.
; APPLICANT: TERRAS, FRANKY R.G.
; APPLICANT: VANDERLEYDEN, JOZEF
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

us-09-910-582b-10.ra1

Wed Jul 16 13:41:33 2003

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/377,687
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION NUMBER: US 08/002,480
; FILING DATE: 04-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-377-687-24

Query Match 51.4%; Score 38; DB 1; Length 49;
Best Local Similarity 41.7%; Pred. No. 18;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CRSWNKADNRSC 12
Db 24 CRNWSAKHGAC 35

RESULT 10
US-08-777-192-15
; Sequence 15, Application US/08777192
; Patent No. 5824869
; GENERAL INFORMATION:
; APPLICANT: BROEKAERT, WILLEM F.
; APPLICANT: CAMMUE, BRUNO P.A.
; APPLICANT: OSBORN, RUPERT W.
; APPLICANT: REES, SARAH B.
; APPLICANT: TERRAS, FRANKY R.G.
; APPLICANT: VANDERLEYDEN, JOZEF
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/777,192
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: US 08/002,480
; FILING DATE: 04-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-777-192-15

Query Match 51.4%; Score 38; DB 2; Length 49;
Best Local Similarity 41.7%; Pred. No. 18;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CRSWNKADNRSC 12
Db 24 CRNWSAKHGAC 35

RESULT 9
US-08-656-318A-7
; Sequence 7, Application US/08656318A
; Patent No. 5750504
; GENERAL INFORMATION:
; APPLICANT: BROEKAERT, WILLEM F.
; APPLICANT: CAMMUE, BRUNO P.A.
; APPLICANT: OSBORN, RUPERT W.
; APPLICANT: REES, SARAH B.
; TITLE OF INVENTION: ANTIMICROBIAL PROTEINS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; ADDRESSEE: Intellectual Property Group of
; ADDRESSEE: PILLSBURY MADISON & SUTRO LLP
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,318A
; FILING DATE: 12-JUN-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION NUMBER: GB 9326424.0
; FILING DATE: 24-DEC-1993
; APPLICATION NUMBER: PCT/GB94/02766
; FILING DATE: 19-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 224199/SEE37925/UST
; TELECOMMUNICATION INFORMATION:

```

QY 1 CRSWNKADNRSC 12
Db 24 CRNWSAKHGAC 35

RESULT 11

US-08-777-192-24
; Sequence 24, Application US/08777192
; Patent No. 5824869
; GENERAL INFORMATION:
; APPLICANT: BROEKAERT, WILLEM F.
; APPLICANT: CAMMUE, BRUNO P.A.
; APPLICANT: OSBORN, RUPERT W.
; APPLICANT: REES, SARAH B.
; APPLICANT: TERRAS, FRANKY R.G.
; APPLICANT: VANDERLEYDEN, JOZEF
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/777,192
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/002,480
; FILING DATE: 04-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KORULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-777-192-24

Query Match 51.4%; Score 38; DB 2; Length 49;
Best Local Similarity 41.7%; Pred. No. 18;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CRSWNKADNRSC 12
Db 24 CRNWSAKHGAC 35

RESULT 12

US-08-956-459-7
; Sequence 7, Application US/08956459
; Patent No. 5919918
; GENERAL INFORMATION:
; APPLICANT: BROEKAERT, WILLEM F.
; APPLICANT: CAMMUE, BRUNO P.A.
; APPLICANT: OSBORN, RUPERT W.
; APPLICANT: REES, SARAH B.
; TITLE OF INVENTION: ANTIMICROBIAL PROTEINS

NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY MADISON & SUTRO LLP
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,459
FILING DATE: 22-OCT-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/656,318
FILING DATE: 12-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB94/02766
FILING DATE: 19-DEC-1994
APPLICATION NUMBER: GB 9326424.0
FILING DATE: 24-DEC-1993
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE: CB-AMP1
US-08-956-459-7

Query Match 51.4%; Score 38; DB 2; Length 49;
Best Local Similarity 41.7%; Pred. No. 18;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CRSWNKADNRSC 12
Db 24 CRNWSAKHGAC 35

RESULT 13

US-08-971-982-15
; Sequence 15, Application US/08971982
; Patent No. 6187904
; GENERAL INFORMATION:
; APPLICANT: BROEKAERT, WILLEM F.
; CAMMUE, BRUNO P.A.
; OSBORN, RUPERT W.
; REES, SARAH B.
; TERRAS, FRANKY R.G.
; VANDERLEYDEN, JOZEF
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971,982

;; FILING DATE: 17-NO. 6187904-1997
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/002,480
;; FILING DATE: 04-JAN-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: KOKULIS, PAUL N.
;; REGISTRATION NUMBER: 16,773
;; REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-861-3000
;; TELEFAX: 202-822-0944
;; INFORMATION FOR SEQ ID NO: 15:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 49 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-08-971-982-15

Query Match 51.4%; Score 38; DB 4; Length 49;
Best Local Similarity 41.7%; Pred. No. 18;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CRSWNKADNRSC 12
DB 24 CRNWSAKHGAC 35

RESULT 14

;; Sequence 24, Application US/08971982
;; Patent No. 6187904
;; GENERAL INFORMATION:
;; APPLICANT: BROEKERT, WILLEM F.
;; CAMMUE, BRUNO P.A.
;; OSBORN, RUPERT W.
;; REES, SARAH B.
;; TERRAS, FRANKY R.G.
;; VANDERLEYDEN, JOZEF
;; TITLE OF INVENTION: BIOCIDAL PROTEINS
;; NUMBER OF SEQUENCES: 59
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
;; STREET: 1100 NEW YORK AVENUE, N.W.
;; CITY: WASHINGTON
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20005

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/971,982
;; FILING DATE: 17-NO. 6187904-1997
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/002,480
;; FILING DATE: 04-JAN-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: KOKULIS, PAUL N.
;; REGISTRATION NUMBER: 16,773
;; REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-861-3000
;; TELEFAX: 202-822-0944
;; INFORMATION FOR SEQ ID NO: 24:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 49 amino acids

;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-08-971-982-24

Query Match 51.4%; Score 38; DB 4; Length 49;
Best Local Similarity 41.7%; Pred. No. 18;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CRSWNKADNRSC 12
DB 24 CRNWSAKHGAC 35

RESULT 15

;; US-09-199-637A-277.
;; Sequence 277, Application US/09199637A
;; Patent No. 6355411
;; GENERAL INFORMATION:
;; APPLICANT: Ausubel, Frederick
;; APPLICANT: Goodman, Howard M.
;; APPLICANT: Rahme, Laurence G.
;; APPLICANT: Mahajan-Wikios, Shalina
;; APPLICANT: Tan, Man-Wah
;; APPLICANT: Cao, Hui
;; APPLICANT: Drenkard, Eliana
;; APPLICANT: Tsongalis, John
;; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
;; FILE REFERENCE: 00786/361002
;; CURRENT APPLICATION NUMBER: US/09/199,637A
;; CURRENT FILING DATE: 1998-11-25
;; PRIOR APPLICATION NUMBER: 60/066,517
;; PRIOR FILING DATE: 1997-11-25
;; NUMBER OF SEQ ID NOS: 437
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 277
;; LENGTH: 558
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-277

Query Match 51.4%; Score 38; DB 4; Length 558;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CRSWNKADN 9
DB 532 CRSWKKKKN 540

Search completed: July 16, 2003, 13:17:09
Job time : 15.0488 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:15:24 ; Search time 22.8293 Seconds
(without alignments)
62.425 Million cell updates/sec

Title: US-09-910-582B-10
Perfect score: 74
Sequence: 1 CRSWNKADNRSC 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*

1:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
6:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10:	/cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12:	/cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	74	100.0	12	9	US-09-910-582B-10
2	74	100.0	12	10	US-09-782-650-5
3	41	55.4	170	9	US-10-101-464A-769
4	41	55.4	262	9	US-10-080-170-459
5	41	55.4	998	9	US-10-101-464A-931
6	39	52.7	144	9	US-10-101-464A-789
7	39	52.7	547	9	US-10-101-464A-928
8	38	51.4	15	9	US-09-910-009A-445
9	38	51.4	16	9	US-09-910-009A-415
10	38	51.4	49	10	US-09-759-584-15
11	38	51.4	49	10	US-09-759-584-24
12	38	51.4	49	10	US-09-917-340-60
13	38	51.4	70	9	US-09-910-009A-228
14	38	51.4	71	9	US-09-910-009A-111
15	38	51.4	245	10	US-09-864-761-33581
16	38	51.4	526	9	US-10-152-548-6
17	38	51.4	558	9	US-09-975-719-277
18	37	50.0	165	9	US-10-001-054-44
19	37	50.0	165	9	US-10-227-884-230

20	37	50.0	165	9	US-10-230-163-230	Sequence 230, App
21	37	50.0	165	9	US-10-218-631-230	Sequence 230, App
22	37	50.0	165	9	US-10-230-338-230	Sequence 230, App
23	37	50.0	165	9	US-10-230-414-230	Sequence 230, App
24	37	50.0	165	9	US-10-216-159A-230	Sequence 230, App
25	37	50.0	165	9	US-10-218-849-230	Sequence 230, App
26	37	50.0	165	9	US-10-227-873-230	Sequence 230, App
27	37	50.0	165	9	US-10-227-883-230	Sequence 230, App
28	37	50.0	165	9	US-10-219-076-230	Sequence 230, App
29	37	50.0	165	9	US-10-230-434-230	Sequence 230, App
30	37	50.0	165	9	US-10-219-003-230	Sequence 230, App
31	37	50.0	165	9	US-10-219-075-230	Sequence 230, App
32	37	50.0	165	9	US-10-219-464-230	Sequence 230, App
33	37	50.0	165	9	US-10-219-466-230	Sequence 230, App
34	37	50.0	165	9	US-10-219-479-230	Sequence 230, App
35	37	50.0	165	9	US-10-219-481-230	Sequence 230, App
36	37	50.0	165	9	US-10-230-260-230	Sequence 230, App
37	37	50.0	165	9	US-10-232-231-230	Sequence 230, App
38	37	50.0	165	9	US-10-232-233-230	Sequence 230, App
39	37	50.0	165	9	US-10-216-165-230	Sequence 230, App
40	37	50.0	165	9	US-10-218-956-230	Sequence 230, App
41	37	50.0	165	9	US-10-219-468-230	Sequence 230, App
42	37	50.0	165	9	US-10-219-478-230	Sequence 230, App
43	37	50.0	165	9	US-10-219-536-230	Sequence 230, App
44	37	50.0	165	9	US-10-219-072-230	Sequence 230, App
45	37	50.0	165	9	US-10-219-470-230	Sequence 230, App

ALIGNMENTS

RESULT 1
US-09-910-582B-10
; Sequence 10, Application US/09910582B
; Publication No. US20030045476A1
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Mackenna, Deirdre A.
; TITLE OF INVENTION: Heart Homing Conjugates
; FILE REFERENCE: P-LJ 4857
; CURRENT APPLICATION NUMBER: US/09/910, 582B
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 09/326,718
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-09-910-582B-10

Query Match 100.0%; Score 74; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CRSWNKADNRSC 12
| | | | | | | | | | | | | |
Db 1 CRSWNKADNRSC 12

RESULT 2
US-09-782-650-5
; Sequence 5, Application US/09782650
; Patent No. US20020019350A1
; GENERAL INFORMATION:
; APPLICANT: Levine, Arnold J.
; APPLICANT: Mitterex, Artur
; APPLICANT: Falkner, Falko-Guenter
; APPLICANT: Schefflinger, Friedrich
; APPLICANT: Dörner, Friedrich

us-09-910-582b-10.rapb

Wed Jul 16 13:41:33 2003

```

; APPLICANT: Edwards Lifesciences Corporation
; TITLE OF INVENTION: Targeted Angiogenesis
; FILE REFERENCE: 20553D-000611US
; CURRENT APPLICATION NUMBER: US/09/782,650
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: US 09/324,079
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 09/327,045
; PRIOR FILING DATE: 1999-06-07
; PRIOR APPLICATION NUMBER: PCT/US00/14988
; PRIOR FILING DATE: 2000-05-31
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:targeting
; OTHER INFORMATION: molecule
US-09-782-650-5

Query Match          100.0%; Score 74; DB 10; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CRSWNKADNRSC 12
Db      1 CRSWNKADNRSC 12
      |||||
RESULT 3
US-10-101-464A-769
; Sequence 769, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 769
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-769

Query Match          55.4%; Score 41; DB 9; Length 170;
Best Local Similarity 54.5%; Pred. No. 32;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      2 RSNWKADNRSC 12
Db      48 RWNENDNSPC 58
      |::|:|:|
RESULT 4
US-10-080-170-459
; Sequence 459, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 459
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-459

Query Match          55.4%; Score 41; DB 9; Length 262;
Best Local Similarity 70.0%; Pred. No. 48;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 RSNWKADNRSC 11
Db      7 RSWKRFDRS 16
      |||:|:|
RESULT 5
US-10-101-464A-931
; Sequence 931, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 931
; LENGTH: 998
; TYPE: PRT
; ORGANISM: Pinus radiata
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(998)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-101-464A-931

Query Match          55.4%; Score 41; DB 9; Length 998;
Best Local Similarity 54.5%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      2 RSNWKADNRSC 12
Db      48 RWNENDNSPC 58
      |:|:|:|:|
RESULT 6
US-10-101-464A-789
; Sequence 789, Application US/10101464A
; Publication No. US20030046728A1
```

GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 789
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-789

Query Match 52.7%; Score 39; DB 9; Length 144;
Best Local Similarity 60.0%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 SWNKADNRSC 12
| | | | |
Db 50 SWNEEDNSPC 59

RESULT 7

US-10-101-464A-928
; Sequence 928, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 928
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-928

Query Match 52.7%; Score 39; DB 9; Length 547;
Best Local Similarity 60.0%; Pred. No. 2e-02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 SWNKADNRSC 12
| | | | |
Db 50 SWNEEDNSPC 59

RESULT 8

US-09-910-009A-445
; Sequence 445, Application US/09910009A
; Publication No. US20030050234A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Watkins, Maren
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Shon, Ki-Joon
; APPLICANT: Jacobsen, Richard
; APPLICANT: Jones, Robert M.
; APPLICANT: Cartier, G. Edward
; APPLICANT: Shen, Greg S.
; APPLICANT: Wagstaff, John D.
; TITLE OF INVENTION: Mu-Conopeptides
; FILE REFERENCE: 2314-242
; CURRENT APPLICATION NUMBER: US/09/910,009A
; CURRENT FILING DATE: 2001-07-23
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: US 60/219,619
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/264,319
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: US 60/277,270
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 520
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 445
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Conus textile
US-09-910-009A-445

Query Match 51.4%; Score 38; DB 9; Length 15;
Best Local Similarity 50.0%; Pred. No. 9.8;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CRSNKADNRSC 12
| | | | |
Db 1 CCSWDVCDHPSC 12

RESULT 9

US-09-910-009A-415
; Sequence 415, Application US/09910009A
; Publication No. US20030050234A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Watkins, Maren
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Shon, Ki-Joon
; APPLICANT: Jacobsen, Richard
; APPLICANT: Jones, Robert M.
; APPLICANT: Cartier, G. Edward
; APPLICANT: Shen, Greg S.
; APPLICANT: Wagstaff, John D.
; TITLE OF INVENTION: Mu-Conopeptides
; FILE REFERENCE: 2314-242
; CURRENT APPLICATION NUMBER: US/09/910,009A
; CURRENT FILING DATE: 2001-07-23
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: US 60/219,619
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/264,319

Qy 3 SWNKADNRSC 12
| | | | |
Db 50 SWNEEDNSPC 59

;; PRIOR FILING DATE: 2001-01-29
;; PRIOR APPLICATION NUMBER: US 60/277,270
;; PRIOR FILING DATE: 2001-03-21
;; NUMBER OF SEQ ID NOS: 520
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 415
;; LENGTH: 16
;; TYPE: PRT
;; ORGANISM: Conus gloriamaris
US-09-910-009A-415

Query Match 51.4%; Score 38; DB 9; Length 16;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CRSWNKADNRSC 12
DB 1 CCSWDVCDHPSC 12

RESULT 10
US-09-759-584-15
; Sequence 15, Application US/09759584
; Patent No. US20010014732A1
; GENERAL INFORMATION:

;; APPLICANT: BROEKAERT, WILLEM F.
;; APPLICANT: CAMMUE, BRUNO P.A.
;; APPLICANT: OSBORN, RUPERT W.
;; APPLICANT: REES, SARAH B.
;; APPLICANT: TERRAS, FRANKY R.G.
;; APPLICANT: VANDERLEYDEN, JOZEF
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARB & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/759,584
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/377,687
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-759-584-15

Query Match 51.4%; Score 38; DB 10; Length 49;
Best Local Similarity 41.7%; Pred. No. 30;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CRSWNKADNRSC 12

Db 24 CRNWESAKHGAC 35

RESULT 11
US-09-759-584-24
; Sequence 24, Application US/09759584
; Patent No. US20010014732A1
; GENERAL INFORMATION:
;; APPLICANT: BROEKAERT, WILLEM F.
;; APPLICANT: CAMMUE, BRUNO P.A.
;; APPLICANT: OSBORN, RUPERT W.
;; APPLICANT: REES, SARAH B.
;; APPLICANT: TERRAS, FRANKY R.G.
;; APPLICANT: VANDERLEYDEN, JOZEF
; TITLE OF INVENTION: BIOCIDAL PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARB & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/759,584
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/377,687
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 99042/SEE.36525/US/A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-759-584-24

Query Match 51.4%; Score 38; DB 10; Length 49;
Best Local Similarity 41.7%; Pred. No. 30;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CRSWNKADNRSC 12
DB 24 CRNWESAKHGAC 35

RESULT 12
US-09-917-340-60
; Sequence 60, Application US/09917340
; Patent No. US20020090369A1
; GENERAL INFORMATION:
;; APPLICANT: Murphy, Christopher J.
;; APPLICANT: McAnulty, Jonathan F.
;; APPLICANT: Reid, Ted W.
; TITLE OF INVENTION: Transplant Media
; FILE REFERENCE: TPLANT-06468
; CURRENT APPLICATION NUMBER: US/09/917,340
; CURRENT FILING DATE: 2001-07-29

; PRIOR APPLICATION NUMBER: 60/221,632
 ; PRIOR FILING DATE: 2000-07-28
 ; PRIOR APPLICATION NUMBER: 60/249,602
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/290,932
 ; PRIOR FILING DATE: 2001-05-15
 ; NUMBER OF SEQ ID NOS: 96
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 60
 ; LENGTH: 49
 ; TYPE: PRT
 ; ORGANISM: Clitoria ternatea
 ; US-09-917-340-60

Query Match 51.4%; Score 38; DB 10; Length 49;
 Best Local Similarity 41.7%; Pred. No. 30;
 Matches 5; Conservative 3; Mismatches 4; Indels 4; Gaps 0;

QY 1 CRSWKADNRSC 12
 DB 24 CRWESAKHGAC 35

RESULT 13

; US-09-910-009A-228
 ; Sequence 228, Application US/09910009A
 ; Publication No. US20030050234A1

; GENERAL INFORMATION:
 ; APPLICANT: University of Utah Research Foundation
 ; APPLICANT: Cognetix, Inc.
 ; APPLICANT: Olivera, Baldomero M.
 ; APPLICANT: McIntosh, J. Michael
 ; APPLICANT: Garrett, James E.
 ; APPLICANT: Watkins, Maren
 ; APPLICANT: Cruz, Lourdes J.
 ; APPLICANT: Shon, Ki-Joon
 ; APPLICANT: Jacobsen, Richard
 ; APPLICANT: Jones, Robert M.
 ; APPLICANT: Cartier, G. Edward
 ; APPLICANT: Shen, Greg S.
 ; APPLICANT: Wagstaff, John D.
 ; TITLE OF INVENTION: Mu-Conopeptides
 ; FILE REFERENCE: 2314-242
 ; CURRENT APPLICATION NUMBER: US/09/910,009A
 ; PRIOR FILING DATE: 2001-07-23
 ; PRIOR APPLICATION NUMBER: US 60/219,619
 ; PRIOR FILING DATE: 2000-07-21
 ; PRIOR APPLICATION NUMBER: US 60/245,157
 ; PRIOR FILING DATE: 2000-11-03
 ; PRIOR APPLICATION NUMBER: US 60/264,319
 ; PRIOR FILING DATE: 2001-01-29
 ; PRIOR APPLICATION NUMBER: US 60/277,270
 ; PRIOR FILING DATE: 2001-03-21
 ; NUMBER OF SEQ ID NOS: 520
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 228
 ; LENGTH: 70
 ; TYPE: PRT
 ; ORGANISM: Conus textile
 ; US-09-910-009A-228

Query Match 51.4%; Score 38; DB 9; Length 70;
 Best Local Similarity 50.0%; Pred. No. 41;
 Matches 6; Conservative 2; Mismatches 4; Indels 4; Gaps 0;

QY 1 CRSWKADNRSC 12
 DB 55 CCSWDVCDHPSC 66

RESULT 14

; US-09-910-009A-111
 ; Sequence 111, Application US/09910009A

; Publication No. US20030050234A1
 ; GENERAL INFORMATION:
 ; APPLICANT: University of Utah Research Foundation
 ; APPLICANT: Cognetix, Inc.
 ; APPLICANT: Olivera, Baldomero M.
 ; APPLICANT: McIntosh, J. Michael
 ; APPLICANT: Garrett, James E.
 ; APPLICANT: Watkins, Maren
 ; APPLICANT: Cruz, Lourdes J.
 ; APPLICANT: Shon, Ki-Joon
 ; APPLICANT: Jacobsen, Richard
 ; APPLICANT: Jones, Robert M.
 ; APPLICANT: Cartier, G. Edward
 ; APPLICANT: Shen, Greg S.
 ; APPLICANT: Wagstaff, John D.
 ; TITLE OF INVENTION: Mu-Conopeptides
 ; FILE REFERENCE: 2314-242
 ; CURRENT APPLICATION NUMBER: US/09/910,009A
 ; PRIOR FILING DATE: 2001-07-23
 ; PRIOR APPLICATION NUMBER: US 60/219,619
 ; PRIOR FILING DATE: 2000-07-21
 ; PRIOR APPLICATION NUMBER: US 60/245,157
 ; PRIOR FILING DATE: 2000-11-03
 ; PRIOR APPLICATION NUMBER: US 60/264,319
 ; PRIOR FILING DATE: 2001-01-29
 ; PRIOR APPLICATION NUMBER: US 60/277,270
 ; PRIOR FILING DATE: 2001-03-21
 ; NUMBER OF SEQ ID NOS: 520
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 111
 ; LENGTH: 71
 ; TYPE: PRT
 ; ORGANISM: Conus gloriamaris
 ; US-09-910-009A-111

Query Match 51.4%; Score 38; DB 9; Length 71;
 Best Local Similarity 50.0%; Pred. No. 42;
 Matches 6; Conservative 2; Mismatches 4; Indels 4; Gaps 0;

QY 1 CRSWKADNRSC 12
 DB 55 CCSWDVCDHPSC 66

RESULT 15

; US-09-864-761-33581
 ; Sequence 33581, Application US/09864761
 ; Patent No. US20020048763A1

; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aeomica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664

Wed Jul 16 13:41:33 2003

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33581
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF000133.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EST_HUMAN HIT: A1796739.1, EVALUATE 5.00e-44
; OTHER INFORMATION: SWISSPROT HIT: Q57814, EVALUATE 2.70e+00
; OTHER INFORMATION: EST_HUMAN HIT: AV716401.1, EVALUATE 4.00e-48
US-09-864-761-33581
```

```
Query Match          51.4%; Score 38; DB 10; Length 245;
Best Local Similarity 41.7%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 CRSWKNADNRSC 12
Db      53 CTSWNRSTKLC 64
```

Search completed: July 16, 2003, 13:36:00
Job time : 23.8293 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:09:18 ; Search time 15.8049 Seconds
(without alignments)
72.991 Million cell updates/sec

Title: US-09-910-582B-10
Perfect score: 74
Sequence: 1 CRSWNKADNRSC 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	43.5	58.8	760	2	D71445	hypothetical prote
2	42	56.8	417	2	T20327	hypothetical prote
3	42	56.8	2347	1	TVHURS	kinase-related pro
4	41	55.4	262	2	H70953	hypothetical prote
5	41	55.4	1162	2	H7852	hypothetical prote
6	41	55.4	1169	2	T71639	hypothetical prote
7	41	55.4	2340	2	I48310	kinase-related pro
8	40	54.1	160	2	C65057	hypothetical prote
9	40	54.1	406	2	T30141	hypothetical prote
10	40	54.1	497	2	T23132	hypothetical prote
11	40	54.1	615	2	T20187	hypothetical prote
12	39	52.7	50	2	S6218	defensin AMPl - Ae
13	39	52.7	157	2	D4860	hypothetical prote
14	39	52.7	286	2	A86665	outer membrane lip
15	39	52.7	344	2	T20209	hypothetical prote
16	39	52.7	2338	2	I73957	kinase-related pro
17	38	51.4	49	2	S66219	defensin AMPl - Cl
18	38	51.4	234	2	H89875	hypothetical prote
19	38	51.4	235	2	B68897	hypothetical prote
20	38	51.4	259	2	T27783	hypothetical prote
21	38	51.4	479	2	G36501	probable serine ca
22	38	51.4	482	1	ERAD24	early ElB 54K prot
23	38	51.4	491	2	F96818	hypothetical prote
24	38	51.4	512	2	A05196	hypothetical prote
25	38	51.4	527	2	T27572	hypothetical prote
26	38	51.4	1083	2	T23031	hypothetical prote
27	37.5	50.7	209	2	T05562	hypothetical prote
28	37.5	50.7	856	1	A45394	env polyprotein pr
29	37	50.0	64	2	H70210	hypothetical prote

ALIGNMENTS

RESULT 1

D71445
hypothetical protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
C:Accession: D71445
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; G.
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; G.
anavath, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomen
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.;
C.; Chludzinski, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis t
A:Reference number: A71400; MUID:98121113; PMID:9461215
A:Accession: D71445
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-760 <BEV>
A:Cross-references: GB:297343; NID:g2245073; PID:g327545; PID:g2245115
C:Genetics:
A:Map position: 4COP9-4G3845

Query Match	58.8%	Score 43.5;	DB 2;	Length 760;
Best Local Similarity	66.7%;	Pred. No. 17;		
Matches 8;	Conservative	1;	Mismatches 2;	Indels 1; Gaps 1;
QY	1	CRSWNKADNRSC	12	
Db	411	CRSWNSA-SRGC	421	

RESULT 2

T20327
 Z
 hypothetical protein D1086.6 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T20327
 R:Smeye, R.
 submitted to the EMBL Data Library, November 1996
 A:Reference number: Z19257
 A:Accession: T20327
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1'417 <NIL>
 A:Cross-references: EMBL:Z81491; PIDN:CAB04023.1; GSPDB:GN00023; CESP:D1086.6
 A:Experimental source: clone D1086
 C:Genetics:
 A:Gene: CESP:D1086.6
 A:Map position: 5
 A:Introns: 40/3; 171/3; 207/3; 222/1; 303/1; 366/3; 402/3

Wed Jul 16 13:41:33 2003

Query Match 56.8%; Score 42; DB 1; Length 2347;
Best Local Similarity 41.7%; Pred. No. 86;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CRSWNKADNRSC 12
| | | | |
Db 61 CHEWNSVDQKNC 72

RESULT 4

H70953
hypothetical protein RV1249c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: H70953
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: H70953
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-262 <COL>
A:Cross-references: GB:AL021006; GB:AL123456; NID:g3242277; PIDN:CAAL15905.1; PID:el21
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV1249c

Query Match 55.4%; Score 41; DB 2; Length 262;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RSNWKADNRS 11
| | | | |
Db 7 RSMKRFDRNS 16

RESULT 5

B97852
hypothetical protein RC1218 [imported] - Rickettsia conorii (strain Malish 7)
C:Species: Rickettsia conorii
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
C:Accession: B97852
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.;
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: B97852
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1162 <KUR>
A:Cross-references: GB:AE006914; PIDN:AAL03756.1; PID:gl5620350; GSPDB:GN00173
C:Genetics:
A:Gene: RC1218
C:Superfamily: Rickettsia prowazekii hypothetical protein RP785

Query Match 55.4%; Score 41; DB 2; Length 1162;
Best Local Similarity 54.5%; Pred. No. 66;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CRSWNKADNRS 11
| | | | |
Db 65 CPPWNKSDGRN 75

RESULT 6

C71639
hypothetical protein RP785 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000

Query Match 56.8%; Score 42; DB 2; Length 417;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

QY 1 CRSWNKADNR--RSC 12
| : | | | | : |
Db 36 CETWNKCDVVKRC 49

RESULT 3

TVHURS
N:Alternate names: protein-tyrosine kinase mcf3 (activated ros-1)
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) ros-1
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1988 #sequence_revision 07-Oct-1994 #text_change 11-Jun-1999
C:Accession: A35512; A25223; A24421; A33081
R:Birchmeier, C.; O'Neill, K.; Riggs, M.; Wigler, M.
Proc. Natl. Acad. Sci. U.S.A. 87, 4799-4803, 1990
A:Title: Characterization of ROS1 cDNA from a human glioblastoma cell line.
A:Reference number: A35512; MUID:90280463; PMID:2352949
A:Accession: A35512
A:Molecule type: mRNA
A:Residues: 1-2212, N', 2214-2227, 'QC', 2229-2347 <BIR>
A:Cross-references: GB:M34353
A:Experimental source: glioblastoma cell line SW-1088
R:Matsumine, H.; Wang, L.H.; Shibuya, M.
Mol. Cell. Biol. 6, 3000-3004, 1986
A:Title: Human c-ros-1 gene homologous to the v-ros sequence of UR2 sarcoma virus encode
A:Reference number: A25223; MUID:87064611; PMID:3023956
A:Accession: A25223
A:Molecule type: DNA
A:Residues: 1790-2245, 'KFDSEFSFRCFTVN', <MA2>
A:Cross-references: GB:M13368
A:Experimental source: placenta
A:Note: the differences after residue 2245 result from the authors' misinterpretation of
R:Birchmeier, C.; Birnbaum, D.; Waite, G.; Fasano, O.; Wigler, M.
Mol. Cell. Biol. 6, 3109-3116, 1986
A:Title: Characterization of an activated human ros gene.
A:Reference number: A24421; MUID:87064625; PMID:3785223
A:Accession: A24421
A:Molecule type: mRNA
A:Residues: 1854-2261, A', 2263-2347 <BI2>
A:Cross-references: GB:M13980; NID:g337482; PIDN:AAA36580.1; PID:g337483
A:Experimental source: tumor cells
A:Note: the mcf3 oncogene was formed by DNA rearrangement involving fusion of at least
C:Genetics:
A:Gene: GDB:ROS1

A:Map position: 6q22-6q22
A:Introns: 1853/1, 1881/1, 1926/2, 1980/3, 2002/2, 2045/3, 2078/2, 2145/2, 2190/2
C:Superfamily: Kinase-related protein ros; LDL receptor WYTD-containing repeat homology;
F:1-36/Domain: signal sequence #status predicted <SIG>
F:37-2347/Product: kinase-related protein ROS1 #status predicted <MAT>
F:37-1859/Domain: extracellular #status predicted <EXT>
F:335-378/Domain: LDL receptor WYTD-containing repeat homology <YWL>
F:466-503/Domain: LDL receptor WYTD-containing repeat homology <YWA>
F:715-757/Domain: LDL receptor WYTD-containing repeat homology <YWD>
F:758-798/Domain: LDL receptor WYTD-containing repeat homology <YW3>
F:799-838/Domain: LDL receptor WYTD-containing repeat homology <YW4>
F:843-888/Domain: LDL receptor WYTD-containing repeat homology <YW5>
F:893-933/Domain: LDL receptor WYTD-containing repeat homology <YW6>
F:1532-1574/Domain: LDL receptor WYTD-containing repeat homology <YW7>
F:1860-1883/Domain: transmembrane #status predicted <TMN>
F:1884-2347/Domain: intracellular #status predicted <INT>
F:1943-2222/Domain: protein kinase #status predicted <KIN>
F:1951-1959/Region: protein kinase ATP-binding motif
F:52, 114, 123, 324, 352, 471, 607, 628, 706, 714, 732, 939, 961, 1015, 1087, 1090, 1211, 1272, 1330, 1458,
F:1980/Active site: Lys #status predicted
F:2110, 2114, 2115/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #stat

C:Accession: C71639
 R.Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U.
 Nature 396, 133-140, 1998
 A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
 A:Reference number: A71630; MUID:99039499; PMID:9823893
 A:Accession: C71639
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1169 <AND>
 A:Cross-references: GB:AJ235273; GB:AJ235269; NID:g3861237; PIDN:CAA15211.1; PID:g386131
 A:Experimental source: strain Madrid E
 C:Genetics:
 A:Gene: RP785
 C:Superfamily: Rickettsia prowazekii hypothetical protein RP785

Query Match 55.4%; Score 41; DB 2; Length 1169;
 Best Local Similarity 54.5%; Pred. No. 66;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 CRSWNKADNRS 11
 I I I I I I I
 Db 64 CPPWNKSDGRN 74

RESULT 7

I48310
 kinase-related protein ros - mouse
 N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
 C:Accession: I48310; PC1125
 R:Riettmacher, D.; Langholz, O.; Godecke, S.; Sachs, M.; Birchmeier, C.
 Oncogene 9, 3617-3626, 1994
 A:Title: Biochemical and functional characterization of the murine ros protooncogene.
 A:Reference number: I48310; MUID:95060820; PMID:7970722
 A:Accession: I48310
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: mRNA
 A:Residues: 1-2340 <RES>
 A:Cross-references: EMBL:X81650; NID:g547531; PIDN:CAA57310.1; PID:g547532
 R:Narayana, L.; Nagarajan, L.
 Gene 118, 297-298, 1992
 A:Title: A mouse c-ros genomic clone: identification of a highly conserved 22-amino acid
 A:Reference number: PC1125; MUID:92380520; PMID:1324876
 A:Accession: PC1125
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1874-1886, 'K', 1888-1918, 'Q', 1920-1973 <NAR>
 A:Cross-references: GB:M88609; GB:M88610
 A:Note: the authors translated the codon CAA for residue 1919 as Glu, CGG for residue 19
 C:Genetics:
 A:Gene: c-ros
 A:Introns: 33/3; 1855/1; 1876/1
 C:Superfamily: kinase-related protein ros; LDL receptor YWTD-containing repeat homology;
 C:Keywords: ATP; autophosphorylation; duplication; magnesium; phosphoprotein; phosphotra
 F:754-794/Domain: LDL receptor YWTD-containing repeat homology <YW3>
 F:1874-1878/Domain: transmembrane (fragment) #status predicted <TM>
 F:1936-2215/Domain: protein kinase homology <KIN>
 F:1944-1952/Region: protein kinase ATP-binding motif
 F:1973/Active site: Lys #status predicted

Query Match 55.4%; Score 41; DB 2; Length 2340;
 Best Local Similarity 41.7%; Pred. No. 1.3e+02;
 Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CRSWNKADNRS 12
 I I I I I I I
 Db 62 CQFNNSVDQETC 73

RESULT 8

C65057
 hypothetical protein b2759 - Escherichia coli (strain K-12)

C:Species: Escherichia coli
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
 C:Accession: C65057
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: C65057
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-160 <BLAT>
 A:Cross-references: GB:AE000359; GB:U00096; NID:g1789110; PIDN:AA75801.1; PID:g17891
 A:Experimental source: strain K-12, substrain MG1655
 C:Superfamily: Escherichia coli hypothetical protein b2759

Query Match 54.1%; Score 40; DB 2; Length 160;
 Best Local Similarity 54.5%; Pred. No. 16;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RSWNKADNRSC 12
 I I I I I I I
 Db 12 RAWQDNGSC 22

RESULT 9

T30141
 hypothetical protein E02C12.6 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T30141
 R:Pauley, A.; Gattung, S.
 submitted to the EMBL Data Library, December 1995
 A:Description: The sequence of C. elegans cosmid E02C12.
 A:Reference number: Z20743
 A:Accession: T30141
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-406 <PAU>
 A:Cross-references: EMBL:U41995; PIDN:AAA83463.1; CESP:E02C12.6
 C:Genetics:
 A:Gene: CESP:E02C12.6
 A:Introns: 47/3; 72/1; 123/3; 307/3; 337/3

Query Match 54.1%; Score 40; DB 2; Length 406;
 Best Local Similarity 41.7%; Pred. No. 37;
 Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CRSWNKADNRSC 12
 I I I I I I I
 Db 58 CRGWKEASSKVC 69

RESULT 10

T23132
 hypothetical protein H36L18.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T23132
 R:McMurray, A.
 submitted to the EMBL Data Library, May 1997
 A:Reference number: Z19692
 A:Accession: T23132
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-497 <WIL>
 A:Cross-references: EMBL:Z95309; PIDN:CAB08559.1; GSPDB:GN00028; CESP:H36L18.1
 A:Experimental source: clone H36L18
 C:Genetics:
 A:Gene: CESP:H36L18.1
 A:Map position: X
 A:Introns: 8/1; 46/1; 107/1; 152/1; 203/1; 246/3

A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: D64860
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-157 <BLAT>
 A:Cross-references: GB:AE000214; GB:U00096; NID:gl787382; PIDN:AAK74235.1; PID:gl7873
 A:Experimental source: strain K-12, substrain MG1655

Query Match 52.7%; Score 39; DB 2; Length 157;
 Best Local Similarity 41.7%; Pred. No. 22;
 Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CRSWNKADNRSC 12
 Db 142 CQRWPEGRREC 153
 I: | | | | |

RESULT 14

A86665
 outer membrane lipoprotein precursor plpD [imported] - Lactococcus lactis subsp. lact
 C:Species: Lactococcus lactis subsp. lactis
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
 C:Accession: A86665
 R:Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissbach, J.; Eh
 Genome Res. 11, 731-753, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
 A:Reference number: A86625; MUID:21235186; PMID:11337471
 A:Accession: A86665
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-286 <STO>
 A:Cross-references: GB:AE005176; PID:gl2723187; PIDN:AAK04419.1; GSPDB:GN00146
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: plpD
 C:Superfamily: lipoprotein-28

Query Match 52.7%; Score 39; DB 2; Length 286;
 Best Local Similarity 75.0%; Pred. No. 39;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 RSWNKADN 9
 Db 103 KSWNKANN 110
 I: | | | | |

RESULT 15

T20209
 hypothetical protein C54D10.8 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T20209
 R:Dobson, R.
 submitted to the EMBL Data Library, June 1996
 A:Reference number: Z19238
 A:Accession: T20209
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-344 <WIL>
 A:Cross-references: EMBL:Z75531; PIDN:CAA99798.1; GSPDB:GN00023; CESP:C54D10.8
 A:Experimental source: clone C54D10
 C:Genetics:
 A:Gene: CESP:C54D10.8
 A:Map position: 5
 A:Introns: 34/3; 103/3; 129/2; 311/3

Query Match 52.7%; Score 39; DB 2; Length 344;
 Best Local Similarity 66.7%; Pred. No. 47;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 WNKADNRSC 12
 Db 16 WKLDNRIC 24
 I: | | | | |

Query Match 54.1%; Score 40; DB 2; Length 497;
 Best Local Similarity 61.5%; Pred. No. 44;
 Matches 8; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 2 RSWNK--ADNRSC 12
 Db 247 RSWNKENTNREC 259
 I: | | | | |

RESULT 11

T20187
 hypothetical protein C53D6.6 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T20187
 R:Gardner, A.
 submitted to the EMBL Data Library, March 1996
 A:Reference number: Z19235
 A:Accession: T20187
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-615 <WIL>
 A:Cross-references: EMBL:Z70270; PIDN:CAA94225.1; GSPDB:GN00022; CESP:C53D6.6
 A:Experimental source: clone C53D6
 C:Genetics:
 A:Gene: CESP:C53D6.6
 A:Map position: 4
 A:Introns: 44/1; 387/3; 447/3; 528/3

Query Match 54.1%; Score 40; DB 2; Length 615;
 Best Local Similarity 54.5%; Pred. No. 54;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CRSWNKADNRSC 11
 Db 178 CYSWKNKDRKN 188
 I: | | | | |

RESULT 12

S66218
 defensin AMP1 - Aesculus hippocastanum
 N:Alternate names: seed antifungal protein
 C:Species: Aesculus hippocastanum
 C:Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 17-Mar-1999
 C:Accession: S66218
 R:Osborn, R.W.; de Samblanx, G.W.; Thevisen, K.; Goderis, I.; Torrekens, S.; van Leuven
 FEBS Lett. 368, 257-262, 1995
 A:Title: Isolation and characterisation of plant defensins from seeds of Asteraceae, Fab
 A:Reference number: S66218; MUID:95354848; PMID:7628617
 A:Accession: S66218
 A:Molecule type: protein
 A:Residues: 1-50 <OSB>
 C:Keywords: antifungal

Query Match 52.7%; Score 39; DB 2; Length 50;
 Best Local Similarity 41.7%; Pred. No. 7.8;
 Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CRSWNKADNRSC 12
 Db 24 CODWEKASHGAC 35
 I: | | | | |

RESULT 13

D64860
 hypothetical protein bll151 - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
 C:Accession: D64860
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 A: Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.

Search completed: July 16, 2003, 13:16:17
Job time : 17.8049 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:04:48 ; Search time 8.19512 Seconds
(without alignments)
60.733 Million cell updates/sec

Title: US-09-910-582B-10
Perfect score: 74
Sequence: 1 CRSWNKADNRSC 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	42	56.8	2347	1 KROS_HUMAN	P08922 homo sapien
2	41	55.4	1169	1 Y785_RICPR	O05975 rickettsia
3	40	54.1	160	1 YGCK_ECOLI	P76632 escherichia
4	40	54.1	780	1 CUL3_CAEEL	Q17391 caenorhabdi
5	39	52.7	157	1 YMFO_ECOLI	P75980 escherichia
6	38	51.4	234	1 PUR7_STAAM	Q99V31 staphylococ
7	38	51.4	482	1 E1BL_ADE12	P04491 human adeno
8	38	51.4	512	1 ACCD_TOBAC	P12219 nicotiana t
9	37.5	50.7	856	1 ENV_FIVU1	Q03804 feline immu
10	37	50.0	400	1 HCAD_ECOLI	P77650 escherichia
11	37	50.0	843	1 CO7_HUMAN	P10643 homo sapien
12	37	50.0	907	1 LGR5_HUMAN	O75473 homo sapien
13	37	50.0	1106	1 STC_DROME	P40798 drosophila
14	37	50.0	1182	1 HAIR_MOUSE	Q61645 mus musculu
15	37	50.0	1803	1 YJL3_YEAST	P47024 saccharomyc
16	36	48.6	321	1 YC39_CYAPA	P48279 cyanophora
17	36	48.6	638	1 SYN2_ARATH	Q9sw95 arabidopsis
18	36	48.6	869	1 P021_SCICO	Q03279 sciara copr
19	36	48.6	1276	1 GIL1_ENTHI	P32022 entamoeba h
20	36	48.6	1285	1 SL17_ENTHI	P23502 entamoeba h
21	36	48.6	2190	1 CCAD_CHICK	O73700 gallus gall
22	35	47.3	227	1 EXOD_BPT4	P04536 bacterioph
23	35	47.3	289	1 SOV3_CAEEL	P34548 caenorhabdi
24	35	47.3	297	1 YCBI_ECOLI	P75845 escherichia
25	35	47.3	299	1 ALC_RABIT	P01879 oryctolagus
26	35	47.3	436	1 D2DR_MEILGA	O73810 meleagris g
27	35	47.3	463	1 D5DR_FUGRU	P53454 fugu rubrip
28	35	47.3	487	1 DNLI_BPT4	P00970 bacterioph
29	35	47.3	487	1 DNLI_BPT6	P19088 bacterioph
30	35	47.3	775	1 RRPO_TCV	P17460 turnip crin
31	35	47.3	912	1 PGCB_BOVIN	Q28062 bos taurus
32	35	47.3	1255	1 PER2_HUMAN	O15055 homo sapien
33	35	47.3	1610	1 CCAD_MESAU	Q99244 mesocricetu

RESULT 1				
ID	KROS_HUMAN	STANDARD;	PRT;	2347 AA.
AC	P08922; Q15368;			
DT	01-NOV-1988 (Rel. 09, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Proto-oncogene tyrosine-protein kinase ROS precursor (EC 2.7.1.112)			
DE	(c-ros-1).			
GN	ROS1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90280463; PubMed=2352949;			
RA	Birchmeier C., O'Neill K., Riggs M., Wigler M.;			
RT	"Characterization of ROS1 cDNA from a human glioblastoma cell line.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 87:4799-4803(1990).			
RN	[2]			
RP	SEQUENCE OF 1790-2259 FROM N.A.			
RX	MEDLINE=87064611; PubMed=3023956;			
RA	Matsushima H., Wang L.-H., Shibuya M.;			
RT	"Human c-ros-1 gene homologous to the v-ros sequence of UR2 sarcoma			
RT	virus encodes for a transmembrane receptorlike molecule.";			
RL	Mol. Cell. Biol. 6:3000-3004(1986).			
RN	[3]			
RP	SEQUENCE OF 1854-2245 FROM N.A.			
RX	MEDLINE=87064623; PubMed=3785223;			
RA	Birchmeier C., Birnbaum D., Waitches G., Fasano O., Wigler M.;			
RT	"Characterization of an activated human ros gene.";			
RL	Mol. Cell. Biol. 6:3109-3116(1986).			
CC	-1- FUNCTION: THIS IS A PROBABLY A CELL GROWTH OR DIFFERENTIATION			
CC	FACTOR RECEPTOR WITH A TYROSINE-PROTEIN KINASE ACTIVITY.			
CC	-1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein			
CC	tyrosine phosphate.			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN			
CC	RECEPTOR SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collabora-			
CC	tion between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	at the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M34353; AAA60278.1; -			
DR	EMBL; M13599; AAA60277.1; -			
DR	EMBL; M13368; AAA60277.1; JOINED.			
DR	EMBL; M13591; AAA60277.1; JOINED.			
DR	EMBL; M13592; AAA60277.1; JOINED.			
DR	EMBL; M13593; AAA60277.1; JOINED.			
DR	EMBL; M13594; AAA60277.1; JOINED.			

P07149 s fatty aci
Q01668 homo sapien
P27732 rattus norv
P54103 mus musculu
P51449 homo sapien
Q9uac9 mesobuthus
Q9uac8 mesobuthus
P16654 potato viru
P12840 rattus norv
P35010 galdiera s
P35926 bacterioph
Q92an6 listeria in

ALIGNMENTS

```

Query Match          56.98; Score 42; DB 1; Length 2347;
Best Local Similarity 41.7%; Pred. No. 34;
Matches          5; Conservative      2; Mismatches      5; Indels      0; Gaps      0;

QY      1 CRSWNKADNESC 12
      | | | | |
Db      61 CFWNSVDQKNC 72

RESULT 2
Y785_RICPR
ID      Y785_RICPR      STANDARD;      PRT;      1169 AA.
AC      O05975;
DT      30-MAY-2000 (Rel. 39, Created)
DT      30-MAY-2000 (Rel. 39, Last sequence update)
DE      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Hypothetical protein RP785 precursor.
GN      RP785.
OS      Rickettsia prowazekii.
OC      Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC      Rickettsiaceae; Rickettsiae; Rickettsia.
NCBI_TaxID=782;
[1]
RN      SEQUENCE FROM N.A.
RP      STRAIN=Madrid E;
RC      MEDLINE=99039499; PubMed=9823893;
RX      Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA      Sikeritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA      Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT      "The genome sequence of Rickettsia prowazekii and the origin of
RT      mitochondria.";
RL      Nature 396:133-140(1998).
[2]
RN      SEQUENCE OF 159-1169 FROM N.A.
RP      STRAIN=Madrid E;
RC      MEDLINE=97419517; PubMed=9274032;
RX      Andersson J.O., Andersson S.G.E.;
RA      "Genomic rearrangements during evolution of the obligate intracellular
RT      parasite Rickettsia prowazekii as inferred from an analysis of 52015
RT      bp nucleotide sequence.";
RL      Microbiology 143:2783-2795(1997).
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC      or send an email to license@isb-sib.ch).
-----
CC      EMBL; AJ235273; CAAL15211.1; -.
DR      EMBL; Y11779; CAA72464.1; -.
KW      Hypothetical protein; Signal; Complete proteome.
FT      SIGNAL          1      20      POTENTIAL.
FT      CHAIN           21 1169  HYPOTHETICAL PROTEIN RP785.
SQ      SEQUENCE      1169 AA; 129918 MW; E01ACCAE9103E716 CRC64;

Query Match          55.4%; Score 41; DB 1; Length 1169;
Best Local Similarity 54.5%; Pred. No. 25;
Matches          6; Conservative      2; Mismatches      3; Indels      0; Gaps      0;

QY      1 CRSWNKADNRS 11
      | | | | |
Db      64 CPPWNKSDGRN 74

RESULT 3
YGCK_ECOLI
ID      YGCK_ECOLI      STANDARD;      PRT;      160 AA.
AC      P76632; Q46900;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)

```

16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical protein ygck.

DE YGCK OR B2759.
GN Escherichia coli.
OS Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AE000359; AAC75801.1; -;
DR EMBL; U29579; AAA6269.1; ALT_SEQ.
DR Ecogene; EGI3118; ygck.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 160 AA; 18701 MW; F2493AAD46595E5C CRC64;

Query Match 54.1%; Score 40; DB 1; Length 160;
Best Local Similarity 54.5%; Pred. No. 5.3;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 RSNKADNRSC 12
| : | : | : |
DB 12 RANQDLNGSC 22

RESULT 4

ID CUL3_CAEEL STANDARD; PRT; 780 AA.
AC Q17391;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Cul-3 protein.
GN CUL-3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodirinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX MEDLINE-96279828; PubMed-8681378;
RA Kipreos E.T., Lander L.E., Wang J.P., He W.W., Hedgecock E.M.;
RT "cul-1 is required for cell cycle exit in C. elegans and identifies a
RT novel gene family.";
RL Cell 85:829-839 (1996).
CC -1- DEVELOPMENTAL STAGE: HIGHEST LEVELS IN EMBRYOS AND LOWER LEVELS IN
CC LARVAE AND ADULTS.
CC -1- SIMILARITY: BELONGS TO THE CULLIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC

CC EMBL; U58085; AAC47122.1; -;
DR InterPro; IPR001373; Cullin.
DR Pfam; PF00888; Cullin; 1.
DR SMART; SM00182; CULLIN; 1.
DR PROSITE; PS01256; CULLIN_1; 1.
DR PROSITE; PS00069; CULLIN_2; 1.
SQ SEQUENCE 780 AA; 90565 MW; 21212B67BF3136CB CRC64;

Query Match 54.1%; Score 40; DB 1; Length 780;
Best Local Similarity 41.7%; Pred. No. 25;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 CRSNKADNRSC 12
| : | : | : |
DB 219 CKNWLSGDNDAC 230

RESULT 5

ID YMFO_ECOLI STANDARD; PRT; 157 AA.
AC P75980;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ymfo.
GN YMFO OR B1151.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE-97061202; PubMed-8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155 (1996).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AE000214; AAC74235.1; -;
DR EMBL; D90749; BAA35977.1; -;
DR EMBL; D90750; BAA35989.1; -;
DR Ecogene; EGI4253; ymfo.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 157 AA; 17609 MW; 54CFD7BF609ECBA2 CRC64;

Query Match 52.7%; Score 39; DB 1; Length 157;
Best Local Similarity 41.7%; Pred. No. 7.7;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Matches	6:	Conservative	1:	Mismatches	2:	Indels	0:	Gaps	0:
---------	----	--------------	----	------------	----	--------	----	------	----

Qy	1	CRS	WKN	KADN	9
Dh	196	CR	W	KADN	204

```

RESULT 7
EIBL_ADEL2
ID      EIBL_ADEL2      STANDARD;      PRT;      482 AA.
AC      F04491;
DT      13-AUG-1987 (Rel. 05, Created)
DT      13-AUG-1987 (Rel. 05, Last sequence update)
DT      01-JUN-1994 (Rel. 29, Last annotation update)
DE      E1B protein, large T-antigen.
OS      Human adenovirus type 12.
OC      Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX      NCBI_TaxID=28282;
RN      [1]
RN      SEQUENCE FROM N.A.
RP      MEDLINE=82105565; PubMed=6275367;
RX      Kimura T., Sawada Y., Shinawawa M., Shimizu Y., Shiroki K.,
RA      Shimojo H., Sudsaki H., Takanami M., Uemizu Y., Fujinaga K.;
RT      "Nucleotide sequence of the transforming early region E1b of
RT      adenovirus type 12 DNA: structure and gene organization, and
RT      comparison with those of adenovirus type 5 DNA.";
RL      Nucleic Acids Res. 9:6571-6589(1981).
RN      [2]
RN      SEQUENCE FROM N.A.
RP      MEDLINE=82115327; PubMed=7326748;
RX      Bos J.L., Polder L.J., Bernards R., Schrier P.I., van den Elsen P.J.,
RA      van der Eb A.J., van Ormondt H.;
RT      "The 2.2 kb E1b mRNA of human Ad12 and Ad5 codes for two tumor
RT      antigens starting at different AUG triplets.";
RL      Cell 27:121-131(1981).
RN      [3]
RN      SEQUENCE FROM N.A.
RP      MEDLINE=94076430; PubMed=8254750;
RX      Sprengel J., Schmitz B., Heuss-Neitzel D., Zock C., Doerfler W.;
RT      "Nucleotide sequence of human adenovirus type 12 DNA: comparative
RT      functional analysis.";
RL      J. Virol. 68:379-389(1994).

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; V000004; CAA23407.1; -;
 DR EMBL; X73487; CAA51879.1; -;
 DR PIR; A03812; ERA024.
 DR PIR; S33930; S33930.
 DR InterPro; IPR002612; adenoE1B_55kDa.
 DR Pfam; PF01690; Adeno_E1B_55K; 1.
 DR Early protein.
 DR SOF SEQUENCE 482 AA; 53935 MW; 359E8082B7EF3659 CRC64;
 SO SEQUENCE

Query Match	51.4%	Score 38;	DB 1;	Length 482;
Best Local Similarity	60.0%;	Pred. No. 34;		
Matches	6;	Conservative	0;	Mismatches 4;
				Indels 0;
				Gaps 0;

QY	3	SWNKADNRSC	12
ph	259	SWNKVSARGC	268

```

RESULT 8
ACCD_TOBAC
ID ACCD_TOBAC
AC P12219;
STANDARD;      •  PRT;    512 AA.

```

Qy	1	CRS	W	K	A	D	R	S	C	12			
			:		:								
Db	142	C	Q	W	P	E	G	D	R	R	E	C	153

RESULT 6	PUR7_STAM	STANDARD;	PRT;	234 AA.
ID	PUR7_STAM			
AC	Q99V31;			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Phosphoribosylaminimidazole-succinocarboxamide synthase (EC 6.3.2.6)			
DE	(SAICAR synthetase).			
GN	PURC OR SAVI056 OR SA0918 OR MW0949.			
OS	Staphylococcus aureus (strain Mu50 / ATCC 700699),			
OS	Staphylococcus aureus (strain N315), and			
OS	Staphylococcus aureus (strain MW2).			
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.			
OX	NCBI_TaxID=158678, 158679, 196620;			
RP	[1]			
RN	SEQUENCE FROM N.A.			
RC	STRAIN=Mu50 / ATCC 700699, and N315;			
RC	MEDLINE=21311952; PubMed=11418146;			
RX	Kuroda M., Ohta T., Uchiyama I., Saba T., Yuzawa H., Kobayashi I.,			
RX	Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,			
RA	Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,			
RA	Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,			
RA	Sekimizu K., Hirakawa H., Kihara S., Goto S., Yabuzaki J.,			
RA	Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.			
RA	Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;			
RT	"Whole genome sequencing of methicillin-resistant <i>Staphylococcus</i>			
RT	<i>aureus</i> ."			
RL	Lancet 357:1225-1240(2001).			

121] SEQUENCE FROM N. A.
AN STRAIN=MM2;
RC MEDLINE=22040717; PubMed=12044378;
RX Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA";
RL Lancet 359:1819-1827(2002).
CC -1-1- CATALYTIC ACTIVITY: ATP + 1-(5-phosphoribosyl)-4-carboxy-5-
CC aminimidazole + L-aspartate = ADP + phosphate + 1-(5-
CC phosphoribosyl)-4-(N'-succinyl-carboxamide)-5-aminimidazole.
CC -1-1- PATHWAY: De novo purine biosynthesis; seventh step.
CC -1-1- SIMILARITY: BELONGS TO THE SAICAR SYNTHETASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collabora-
CC tion between the Swiss Institute of Bioinformatics and the EMBL outstat-
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is not
CC modified and this statement is not removed. Usage by and for comm-
CC entities requires a license agreement (See <http://www.isb-sib.ch/anno->
CC or send an email to license@sib-sib.ch).

```

CC
DR DR ENBL; APO03361; BAB57228.1; -
DR DR ENBL; APO03132; BAB42163.1; -
DR DR ENBL; APO04825; BAB94814.1; -
DR DR InterPro; IPR001636; SAICAR_synt.
DR DR Pfam; PF01259; SAICAR_synt; 1.
DR DR ProDom; PD003043; SAICAR_synt; 1.
DR DR TIGRFAMS; TIGR00081; purC; 1.
DR DR PROSITE; PS01057; SAICAR_SYNTHETASE_1; FALSE_NEG.
DR DR PROSITE; PS01058; SAICAR_SYNTHETASE_2; 1.
DR DR Purine biosynthesis; Ligase; Complete proteome.
KW SQ SEQUENCE 234 AA; 26962 MW; A38103C9CC04F038 CRC64;

```

Query Match	51.4%;	Score 38;	DB 1;	Length 234;
Best Local Similarity	66.7%;	Pred. No. 17;		

DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Acetyl-coenzyme A carboxylase carboxyl transferase subunit beta
 DE (EC 6.4.1.2) (ACCase beta chain).
 GN ACCD OR YCF11 OR ZFPA.
 OS Nicotiana tabacum (Common tobacco).
 OS Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; eusterids I; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Bright Yellow 4;
 RA Shinzaki K., Ome M., Tanaka M., Wakasugi T., Hayashida N.,
 RA Matsubayashi T., Zaita N., Chunwongse J., Obokata J.,
 RA Yamaguchi-Shinozaki K., Ohto C., Torazawa K., Meng B.-Y., Sugita M.,
 RA Deno H., Kamogashira T., Yamada K., Kusuda J., Takaiwa F., Kato A.,
 RA Tohdoh N., Shimada H., Sugiyama M.;
 RA "The complete nucleotide sequence of the tobacco chloroplast genome:
 RT its gene organization and expression.";
 RL EMBO J. 5:2043-2049(1986).
 CC -1- FUNCTION: PROBABLE SUBUNIT OF A BIOTIN-DEPENDENT CARBOXYLASE
 CC INVOLVED IN CHLOROPLAST CARBON METABOLISM.
 CC -1- CATALYTIC ACTIVITY: ATP + acetyl-CoA + HCO(3)(-) = ADP + phosphate
 CC + malonyl-CoA.
 CC -1- SIMILARITY: BELONGS TO THE ACCD / PCCB FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; 200044; CAA77362.1;
 DR PIR; A05196; A05196.
 DR InterPro; IPR000438; ACoACC.transf.
 DR InterPro; IPR000022; Carboxyl.trans.
 DR Pfam; PF01039; Carboxyl.trans. 1.
 DR PRINTS; PR01070; ACCCTRFRASEB.
 DR TIGRFAMS; TIGR00515; accd; 1.
 KW Fatty acid biosynthesis; Ligase; Zinc-finger; Chloroplast.
 FT ZN_FING 247 269 C4-TYPE (POTENTIAL).
 SQ SEQUENCE 512 AA; 58472 MW; DFB3312E6228868 CRC64;
 Query Match 51.4%; Score 38; DB 1; Length 512;
 Best Local Similarity 60.0%; Pred. No. 36;
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 3 SWNKADNRSC 12
 DB 64 SWNRDNRSSC 73
 RESULT 9
 ENV_FIVU1
 ID ENV_FIVU1 STANDARD; PRT; 856 AA.
 AC Q03804;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE ENV polyprotein precursor (GP150 polyprotein) [Contains: Glycoprotein
 DE GP100; Glycoprotein GP36].
 DE ENV.
 GN Feline immunodeficiency virus (strain UT-113) (FIV).
 OS Viruses; Retroviridae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=36373;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93174954; PubMed=8382405;

RA Verschoor E.J., Hulskotte E.G.J., Ederveen J., Koolen M.J.M.,
 RA Horzinek M.C., Rottier P.J.M.;
 RT "Post-translational processing of the feline immunodeficiency virus
 RT envelope precursor protein.";
 RL Virology 193:433-438(1993).
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X60725; CAA43131.1;
 DR PIR; A45394; A45394.
 DR InterPro; IPR002050; Env_polyprotein.
 DR Pfam; PF00429; Env_polyprotein; 1.
 KW Coat protein; Polyprotein; Glycoprotein; Transmembrane.
 FT CHAIN 1 611 MAJOR GLYCOPROTEIN GP100.
 FT TRANSMEM 612 856 GLYCOPROTEIN GP36.
 FT CARBOHYD 786 812 POTENTIAL.
 FT CARBOHYD 220 220 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 258 258 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 269 269 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 342 342 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 422 422 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 469 469 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 481 481 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 499 499 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 518 518 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 531 531 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 548 548 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 551 551 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 556 556 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 717 717 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 721 721 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 729 729 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 737 737 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 856 AA; 98476 MW; A39DA449DE0BF60C CRC64;
 Query Match 50.7%; Score 37.5; DB 1; Length 856;
 Best Local Similarity 63.6%; Pred. No. 72;
 Matches 7; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
 QY 1 CRSWNKADNRS 11
 DB 462 CR-WNEGDNNS 471
 RESULT 10
 HCAD_ECOLI
 ID HCAD_ECOLI STANDARD; PRT; 400 AA.
 AC P77650; Q08100;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 3-phenylpropionate dioxygenase ferredoxin--NAD(+) reductase component
 DE (EC 1.18.1.3) (Dioxygenin system ferredoxin--NAD(+)) reductase
 DE component).
 GN HCAD OR PHDA OR HCAA4 OR B2542.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.

RC STRAIN-K12 / MCI061;
RX MEDLINE-98269008; PubMed=9603882;
RA Diaz E., Ferrandez A., Garcia J.L.;
RT "Characterization of the hca cluster encoding the dioxigenolytic
pathway for initial catabolism of 3-phenylpropionic acid in
Escherichia coli K-12.";
RL J. Bacteriol. 180:2915-2923(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE-97349980; PubMed=9205837;
RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasubram S.,
Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
K-12 genome corresponding to 50.0-68.8 min on the linkage map and
analysis of its sequence features.";
RL DNA Res. 4:91-113(1997).
CC [1]- FUNCTION: PART OF MULTICOMPONENT 3-PHENYLPROPIONATE DIOXYGENASE.
CC THAT CONVERTS 3-PHENYLPROPIONIC ACID (PP) INTO CIS-3-(3-
CARBOXYETHYL)-3,5-CYCLOHEXADIENE-1,2-DIOL (PP-DIHYDRODIOL).
CC CATALYTIC ACTIVITY: Reduced ferredoxin + NAD(+) = oxidized
ferredoxin + NADH.
CC [1]- COFACTOR: FAD.
CC [1]- PATHWAY: 3-phenylpropionic acid catabolism.
CC [1]- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE
TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (HCAE AND HCAF), A
FERREDOXIN (HCAC) AND A FERREDOXIN REDUCTASE (HCAD).
CC [1]- SIMILARITY: TO OTHER BACTERIAL RING-HYDROXYLATING DIOXYGENASE
FERREDOXIN REDUCTASE COMPONENTS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@lsb-sib.ch).
CC
CC EMBL: Y11070; CAA71952.1;
DR EMBL: AE000340; AAC75595.1;
DR EMBL: D90884; BAA16445.1;
DR EMBL: D90885; BAA16452.1;
DR EMBL: EG13460; hcad.
DR EcoGene: EG13460; hcad.
DR InterPro: IPR001327; FAD_pyr_redox.
DR Pfam: PF00070; pyr_redox; 1.
DR ProDom: PD000139; FAD_pyr_redox; 1.
KW Aromatic hydrocarbons catabolism; Flavoprotein; FAD; NAD;
KW Oxidoreductase; Complete proteome.
FT NP_BIND 5 36 FAD (ADP PART) (POTENTIAL).
FT NP_BIND 146 174 NAD (ADP PART) (POTENTIAL).
SQ SEQUENCE 400 AA; 43978 MW; F5A1A06C4F1DFF36 CRC64;
Query Match 50.08; Score 37; DB 1; Length 400;
Best Local Similarity 45.58; Pred. No. 42;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 CRSWNKADNRS 11
DB 289 CESWENANQA 299

RESULT 11
CO7_HUMAN STANDARD; PRT; 843 AA.
ID CO7_HUMAN
AC P10643; Q92489;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Complement component C7 precursor.
GN C7.
DE Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE-88087145; PubMed=3335508;
RA Discipio R.G., Chakravari D.N., Mueller-Eberhard H.J., Fey G.H.;
RT "The structure of human complement component C7 and the C5b-7
complex.";
RL J. Biol. Chem. 263:549-560(1988).
RN [2]
RP SEQUENCE OF 3-822 FROM N.A.
RX MEDLINE-95248084; PubMed=7730625;
RA Hobart M.J., Fernie B.A., Discipio R.G.;
RT "Structure of the human C7 gene and comparison with the C6, C8A, C8B
and C9 genes.";
RL J. Immunol. 154:5188-5194(1995).
RN [3]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE-20020247; PubMed=10551839;
RA Hofsteenge J., Blommestein M., Hess D., Furmanek A., Miroschnichenko O.;
RT "The four terminal components of the complement system are
C-mannosylated on multiple tryptophan residues.";
RL J. Biol. Chem. 274:32786-32794(1999).
RN [4]
RP VARIANT C7D SER-521.
RX MEDLINE-97025476; PubMed=8871666;
RA Fernie B.A., Wurznner R., Orren A., Morgan B.P., Potter P.C.,
Platonov A.E., Vershinina I.V., Shipulin G.A., Lachmann P.J.,
Hobart M.J.;
RT "Molecular bases of combined subtotal deficiencies of C6 and C7: their
effects in combination with other C6 and C7 deficiencies.";
RL J. Immunol. 157:3648-3657(1996).
RN [5]
RP VARIANT C7D ARG-379.
RX MEDLINE-97361664; PubMed=9218625;
RA Fernie B.A., Orren A., Sheehan G., Schlesinger M., Hobart M.J.;
RT "Molecular bases of C7 deficiency: three different defects.";
RL J. Immunol. 159:1019-1026(1997).
RN [6]
RP VARIANTS C7D GLN-220; GLN-682 AND HIS-687.
RX MEDLINE-99072319; PubMed=9856499;
RA Fernie B.A., Hobart M.J.;
RT "Complement C7 deficiency: seven further molecular defects and their
associated marker haplotypes.";
RL Hum. Genet. 103:513-519(1998).
CC [1]- FUNCTION: C7 IS A CONSTITUENT OF THE MEMBRANE ATTACK COMPLEX.
CC C7 BINDS TO C5B FORMING THE C5B-7 COMPLEX, WHERE IT SERVES
AS A MEMBRANE ANCHOR.
CC [1]- SUBUNIT: MONOMER OR DIMER; AS A C5B-7 COMPLEX IT CAN ALSO
FORM MULTIMERIC ROSETTES.
CC [1]- PTM: C7 HAS 28 DISULFIDE BRIDGES.
CC [1]- DISEASE: Defects in C7 are a cause of component C7 deficiency
(C7D). Patients with C7D suffer from recurrent bacterial
infections, predominantly from Neisseria meningitidis.
CC [1]- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC [1]- SIMILARITY: CONTAINS 2 TSP TYPE-1 DOMAINS.
CC [1]- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC [1]- SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; J03507; AAA51861.1; -
CC EMBL; X86328; CAA60121.1; -
CC EMBL; X86329; CAA60121.1; JOINED.
CC EMBL; X86330; CAA60121.1; JOINED.
CC EMBL; X86331; CAA60121.1; JOINED.
CC EMBL; X86332; CAA60121.1; JOINED.
CC EMBL; X86333; CAA60121.1; JOINED.
CC EMBL; X86334; CAA60121.1; JOINED.
CC EMBL; X86335; CAA60121.1; JOINED.
CC EMBL; X86336; CAA60121.1; JOINED.
CC EMBL; X86337; CAA60121.1; JOINED.
CC EMBL; X86338; CAA60121.1; JOINED.
CC EMBL; X86339; CAA60121.1; JOINED.
CC EMBL; X86340; CAA60121.1; JOINED.
CC EMBL; X86341; CAA60121.1; JOINED.
CC EMBL; X86342; CAA60121.1; JOINED.
CC EMBL; X86343; CAA60121.1; JOINED.
CC EMBL; X86344; CAA60121.1; JOINED.
CC PIR; A27340; A27340.
CC HSSP; Q07954; ICR8.
CC Genew; HGNC:1346; C7.
CC MIM; 217070; -
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR003884; FcR1-MAC.
CC InterPro; IPR002172; LDL-receptor.
CC InterPro; IPR001862; MAC-perforin.
CC InterPro; IPR000436; Sushi_SCR_CCP.
CC InterPro; IPR000884; TSPI.
CC Pfam; PF00057; ldl_recept_a; 2.
CC Pfam; PF00084; sushi; 4.
CC Pfam; PF00090; tsp.1; 4.
CC Pfam; PF01823; MACPF; 2.
CC PRINTS; PR00764; COMPLEMENTC9.
CC SMART; SM00032; CCP; 2.
CC SMART; SM00057; FIMAC; 2.
CC SMART; SM00192; LDLa; 1.
CC SMART; SM00457; MACPF; 1.
CC SMART; SM00209; TSPI; 2.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS01209; LDLRA_1; 1.
CC PROSITE; PS00068; LDLRA_2; 1.
CC PROSITE; PS00279; MAC-PERFORIN; 1.
CC PROSITE; PS00092; TSPI; 2.
KW Complement pathway; Complement alternate pathway; Glycoprotein;
KW Plasma; Membrane attack complex; Cytolysis; Signal; EGF-like domain;
KW Repeat; Sushi; Transmembrane; Disease mutation.
FT SIGNAL 1 22
FT CHAIN 23 843
FT DOMAIN 24 83
FT DOMAIN 83 121
FT TRANSMEM 271 287
FT TRANSMEM 292 311
FT DOMAIN 451 487
FT DOMAIN 497 545
FT DOMAIN 570 627
FT DOMAIN 630 689
FT DOMAIN 695 770
FT DOMAIN 771 843
FT DISULFID 85 96
FT DISULFID 91 109
FT DISULFID 103 119
FT DISULFID 337 353
FT DISULFID 571 613
FT DISULFID 599 626

FT DISULFID 631 673 BY SIMILARITY.
FT DISULFID 659 688 BY SIMILARITY.
FT CARBOHYD 36 36 C-LINKED (MAN).
FT CARBOHYD 202 202 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 503 503 C-LINKED (MAN) (PARTIAL).
FT CARBOHYD 506 506 C-LINKED (MAN) (PARTIAL).
FT CARBOHYD 509 509 C-LINKED (MAN) (PARTIAL).
FT CARBOHYD 754 754 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARIANT 220 220 R -> Q (IN C7D).
FT VARIANT 379 379 /FTID-VAR_012643.
FT VARIANT 521 521 G -> R (IN C7D).
FT VARIANT 521 521 R -> S (IN C7D).
FT VARIANT 682 682 /FTID-VAR_012645.
FT VARIANT 687 687 E -> Q (IN C7D).
FT VARIANT 687 687 /FTID-VAR_012646.
FT CONFLICT 152 152 R -> H (IN C7D).
FT CONFLICT 389 389 /FTID-VAR_012647.
FT CONFLICT 587 587 R -> V (IN REF. 2).
FT CONFLICT 821 822 S -> T (IN REF. 2).
FT CONFLICT 821 822 P -> T (IN REF. 2).
SQ SEQUENCE 843 AA; 93514 MW; DA7817C92E231FA5 CRC64;
Query Match 50.0%; Score 37; DB 1; Length 843;
Best Local Similarity 41.7%; Pred. No. 87;
Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY 1 CRSWNKADNRSC 12
DB 702 CORWEKLQNSRC 713
RESULT 12
LGRS_HUMAN
ID LGRS_HUMAN STANDARD; PRT; 907 AA.
AC 075473; Q9UP75;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 5 precursor
DE (Orphan G protein-coupled receptor HG38) (G protein-coupled receptor
DE 49).
DE GPR49 OR LGR5 OR GPR67.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98308104; PubMed=9642114;
RA McDonald T., Wang R., Bailey W., Xie G., Chen F., Caskey C.T.,
RA Liu Q.;
RT "Identification and cloning of an orphan G protein-coupled receptor of
RT the glycoprotein hormone receptor subfamily.";
RL Biochem. Biophys. Res. Commun. 247:266-270(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=99065210; PubMed=9849958;
RA Hsu S.Y., Liang S.-G., Hsueh A.J.W.;
RT "Characterization of two LGR genes homologous to gonadotropin and
RT thyrotropin receptors with extracellular leucine-rich repeats and a G
RT protein-coupled, seven-transmembrane region.";
RL Mol. Endocrinol. 12:1830-1845(1998).
CC -1- FUNCTION: Orphan receptor. It may be an important receptor for
CC signals controlling growth and differentiation of specific
CC embryonic tissues (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in skeletal muscle, placenta, spinal
CC cord, and various region of brain.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -1- SIMILARITY: CONTAINS 17 LEUCINE-RICH REPEATS (LRR).

CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
DR	EMBL; AF062006; AAC28019.1; ..		
DR	EMBL; AF061444; AAC77911.1; ..		
DR	Genew; HGNC:4504; GPR49.		
DR	MIM; 606667; ..		
DR	HSSP; P23945; 1XUN.		
DR	InterPro; IPR000276; GPCR_Rhodpsn.		
DR	InterPro; IPR001611; LRR.		
DR	InterPro; IPR000372; LRR_Nterm.		
DR	InterPro; IPR000392; LRR_out.		
DR	InterPro; IPR0003591; LRR_typ.		
DR	Pfam; PF000001; 7tm1.1; 1.		
DR	Pfam; PF00560; LRR; 15.		
DR	Pfam; PF01462; LRRT; 1.		
DR	PRINTS; PR00019; LEURICHRPT.		
DR	SMART; SM00370; LRR; 3.		
DR	SMART; SM00013; LRNT; 1.		
DR	SMART; SM00369; LRR_TYP; 8.		
DR	PROSITE; PS00237; G_PROTEIN_RECIP_FL_1; FALSE_NEG.		
DR	PROSITE; PS0262; G_PROTEIN_RECIP_FL_2; 1.		
DR	G-protein coupled receptor; Signal; Transmembrane; Glycoprotein;		
KW	Repeat; Leucine-rich repeat.		
FT	SIGNAL 1 21	POTENTIAL.	
FT	CHAIN 22 907	LEUCINE-RICH REPEAT-CONTAINING G PROTEIN- COUPLED RECEPTOR 5	
FT	DOMAIN 1 561	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM 562 582	1 (POTENTIAL).	
FT	DOMAIN 583 593	CYTOPLASMIC (POTENTIAL).	
FT	TRANSMEM 594 614	2 (POTENTIAL).	
FT	DOMAIN 615 638	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM 639 659	3 (POTENTIAL).	
FT	DOMAIN 660 682	CYTOPLASMIC (POTENTIAL).	
FT	TRANSMEM 683 703	4 (POTENTIAL).	
FT	DOMAIN 704 722	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM 723 743	5 (POTENTIAL).	
FT	DOMAIN 744 767	CYTOPLASMIC (POTENTIAL).	
FT	TRANSMEM 768 788	6 (POTENTIAL).	
FT	DOMAIN 789 802	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM 803 823	7 (POTENTIAL).	
FT	DOMAIN 824 907	CYTOPLASMIC (POTENTIAL).	
FT	REPEAT 64 88	LRR 1.	
FT	REPEAT 89 112	LRR 2.	
FT	REPEAT 114 136	LRR 3.	
FT	REPEAT 137 160	LRR 4.	
FT	REPEAT 162 184	LRR 5.	
FT	REPEAT 185 208	LRR 6.	
FT	REPEAT 209 232	LRR 7.	
FT	REPEAT 233 256	LRR 8.	
FT	REPEAT 257 279	LRR 9.	
FT	REPEAT 281 303	LRR 10.	
FT	REPEAT 304 327	LRR 11.	
FT	REPEAT 328 350	LRR 12.	
FT	REPEAT 351 375	LRR 13.	
FT	REPEAT 377 396	LRR 14.	
FT	REPEAT 397 420	LRR 15.	
FT	REPEAT 422 444	LRR 16.	
FT	REPEAT 445 585	LRR 17.	
FT	REPEAT 564 585	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 63 63	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 77 77	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 208 208	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 500 500	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 792 792	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 90 90	R -> H (IN REF. 2).	
FT	CONFLICT 212 212	L -> W (IN REF. 2).	
FT	CONFLICT 212 212		
FT	SEQUENCE 907 AA; 99997 MW; 822D5C5E6F0D9092 CRC64;		
CC	Query Match 50.0%; Score 37; DB 1; Length 907;		
CC	Best Local Similarity 75.0%; Pred. No. 93;		
CC	Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
QY	4 WNKADNRS 11		
DB	495 WNKGDNS 502		
DR	STC_DROME STANDARD; PRT; 1106 AA.		
AC	P40798;		
DT	01-FEB-1995 (Rel. 31, Created)		
DT	01-FEB-1995 (Rel. 31, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Shuttle craft protein.		
GN	STC.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;		
OC	Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;		
OC	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
ON	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE-Ovary;		
RX	MEDLINE=96104568; PubMed=8524296;		
RA	Stroumbakis N.D., Li Z., Tollas P.P.;		
RT	"A homolog of human transcription factor NF-X1 encoded by the		
RT	Drosophila shuttle craft gene is required in the embryonic central		
RT	nervous system."		
RL	Mol. Cell. Biol. 16:192-201(1996).		
CC	FUNCTION: PLAYS AN ESSENTIAL ROLE DURING THE LATE STAGES OF		
CC	EMBRYONIC NEUROGENESIS. MAY EITHER FINE-TUNE THE GUIDANCE OF		
CC	SPATIAL MAINTENANCE OF THE MIGRATING SNB AND IN NERVE ROOTS,		
CC	WHICH ARE COMPOSED OF AXONS ORIGINATING FROM DISTINCT GROUPS OF		
CC	MOTOR NEURONS AND MAY BE REQUIRED TO EITHER GUIDE OR MAINTAIN THE		
CC	POSITION OF THESE NERVES ALONG A DIRECT AND STRAIGHT PATH TO THEIR		
CC	ULTIMATE TARGETS IN PARTICULAR MUSCLE FIELDS. MAY PLAY A ROLE IN		
CC	EGG CHAMBER DEVELOPMENT AND/OR MAY CONFER ESSENTIAL MATERNAL		
CC	CONTRIBUTIONS TO THE EARLY EMBRYO.		
CC	SUBCELLULAR LOCATION: Nuclear.		
CC	TISSUE SPECIFICITY: OVARIES AND EMBRYONIC CENTRAL NERVOUS SYSTEM.		
CC	DEVELOPMENTAL STAGE: MAJOR EXPRESSION IS SEEN IN THE OVARIES WHILE		
CC	MODERATE LEVELS OF EXPRESSION ARE OBSERVED DURING EMBRYOGENESIS		
CC	AND THROUGHOUT SUBSEQUENT STAGES OF FLY DEVELOPMENT.		
CC	SIMILARITY: TO YEAST YNL023C AND HUMAN NF1.		
CC	SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
CC	EMBL; U09306; AAB60255.1; ..		
DR	TRANSFAC; T01688; ..		
DR	FlyBase; FBgn0001978; stc.		
DR	InterPro; IPR001374; R3H.		
DR	InterPro; IPR000967; znf_NFX1.		
DR	InterPro; IPR001841; znf_ring.		
DR	Pfam; PF01422; zf-NF-X1; 8.		
DR	Pfam; PF01424; R3H; 1.		
DR	SMART; SM00393; R3H; 1.		
DR	SMART; SM00184; RING; 1.		
DR	SMART; SM00438; znf_NFX; 9.		
DR	PROSITE; PS01359; zf_PHD_1; 1.		
DR	PROSITE; PS00518; zf_RING_1; 1.		
DR	PROSITE; PS00089; zf_RING_2; 1.		
DR	Transcription regulation; DNA-binding; Nuclear protein; Repeat;		
KW	Transcription regulation; DNA-binding; Nuclear protein; Repeat;		

KW Alternative splicing; RNA-binding; zinc-finger.
FT DOMAIN 241 268
FT ZN_FING 386 433
FT DOMAIN 464 904
FT REPEAT 464 500
FT REPEAT 517 554
FT REPEAT 575 614
FT REPEAT 634 675
FT REPEAT 723 762
FT REPEAT 834 871
FT REPEAT 875 904
FT VSPPLIC 109 115
SQ SEQUENCE 1106 AA; 123261 MW; 8CE146DB37BB8D65 CRC64;

Query Match 50.0%; Score 37; DB 1; Length 1106;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CR5WNRKDRSC 12
DB 902 CECGRKQNRSC 913

RESULT 14
HAIR_MOUSE STANDARD; PRT; 1182 AA.
AC Q61645;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hairless protein.
GN HR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Skin;
RX MEDLINE=94329587; PubMed=8052649;
RA Cachon-Gonzalez M.B., Fenner S., Coffin J.M., Moran C., Best S.,
Soye J.P.;
RT "Structure and expression of the hairless gene of mice.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:7717-7721(1994).
CC -1- FUNCTION: MAY ACT AS A TRANSCRIPTION FACTOR THAT COULD ACT ON TO
REGULATE ONE OF THE PHASES OF HAIR GROWTH.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN BRAIN, HAIR
FOLLICLES AND INTERFOLLICULAR EPIDERMIS. NO EXPRESSION IN
DERMIS.
CC -1- DISEASE: HR MUTATION PRODUCES A NUMBER OF PLEIOTROPIC EFFECTS
INCLUDING STRUCTURAL ABNORMALITIES OF EPITHELIAL CELLS IN THE HAIR
FOLLICLES, HAIR LOSS TOWARDS THE END OF THE FIRST HAIR GROWTH
CYCLE, AND THE FAILURE OF SUBSEQUENT HAIR GROWTH CYCLES. OLDER
MICE CARRYING AN HR MUTATION HAVE BEEN REPORTED TO POSSESS ALTERED
RATIOS OF T-CELL-DEPENDENT B-CELL RESPONSES. MICE HOMOZYGOUS FOR
HR MUTATION ARE UNIQUELY SENSITIVE TO UV AND CHEMICALLY INDUCED
SKIN TUMORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; Z32675; CAA83587.1; -;
DR MGD; MGI:96223; hr
DR InterPro; IPR003347; TF_JmJc.
DR Pfam; PF02373; JmJc; 1.
KW Zinc-finger; DNA-binding; Nuclear protein; Transcription regulation;
Metal-binding.

FT ZN_FING 595 620 C6-TYPE.
FT DOMAIN 535 540 POLY-GLY.
SQ SEQUENCE 1182 AA; 127182 MW; EFB0EE62AE81B40 CRC64;

Query Match 50.0%; Score 37; DB 1; Length 1182;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 NKADNRSC 12
DB 361 NKADSRAC 368

RESULT 15
YJL3_YEAST STANDARD; PRT; 1803 AA.
AC P47024; P87192;
DT 01-FEB-1996 (Rel. 33, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Transposon Ty4 207.7 kDa hypothetical protein.
GN TY4B OR YJL113W OR J0780.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=97103775; PubMed=8948101;
RA Cziepluch C., Kordes E., Pujol A., Jauniaux J.-C.;
RT "Sequencing analysis of a 40.2 kb fragment of yeast chromosome X
reveals 19 open reading frames including URA2 (5' end), TRK1, PBS2,
SP10, GDL1, RPE1, PHO86, NCA3, ASF1, CCF7, GZF3, two tRNA genes,
three remnant delta elements and a Ty4 transposon.";
RL Yeast 12:1471-1474(1996).

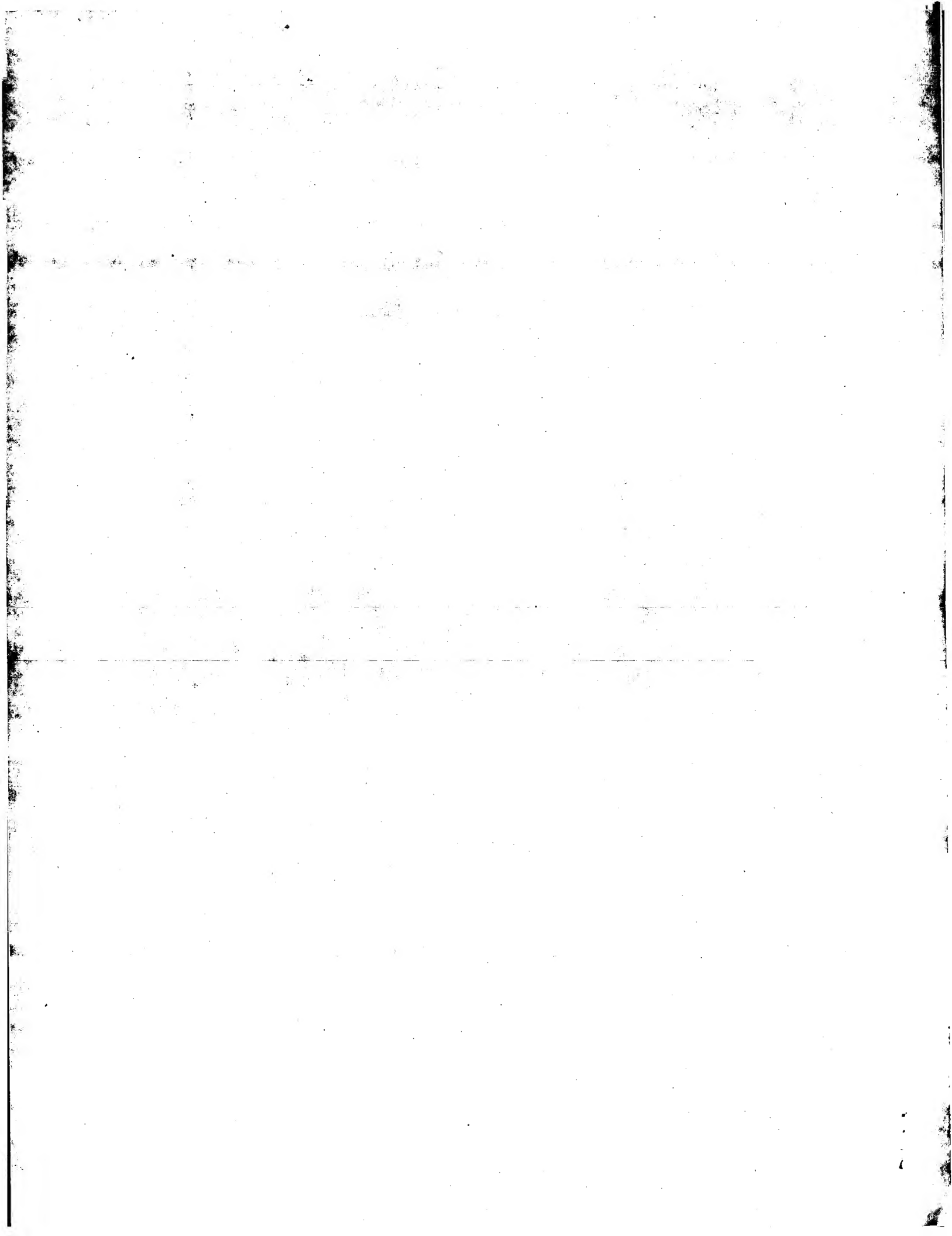
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; Z49389; CAA89409.1; -;
DR SGD; S0003649; YJL113W.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00665; rve; 1.
DR SMART; SM00343; Znf_C2HC; 1.
KW Transposable element; Hypothetical protein.
SQ SEQUENCE 1803 AA; 207693 MW; 16DCD7284A8D52D3 CRC64;

Query Match 50.0%; Score 37; DB 1; Length 1803;
Best Local Similarity 60.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 RSNWKADNR 11
DB 393 KSNWKQKRS 402

Search completed: July 16, 2003, 13:13:16
Job time : 9.19512 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2003, 13:07:18 ; Search time 32.4878 Seconds
(*without alignments)
76.107 Million cell updates/sec

Title: US-09-910-582B-10
Perfect score: 74
Sequence: 1 CRSWNKADNRSC 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriopl.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	59.5	425	2 P70860	P70860 borrelia bu
2	43.5	58.8	760	10 O23598	O23598 arabidopsis
3	41.5	56.1	445	2 O9AMT8	O9AMT8 bradyrhizob
4	41	55.4	262	16 O50464	O50464 mycobacteri
5	41	55.4	477	5 O17729	O17729 caenorhabdi
6	41	55.4	1162	16 O92GA5	O92GA5 rickettsia
7	41	55.4	2340	11 O64736	O64736 mus musculu
8	41	55.4	2340	11 O60705	O60705 mus musculu
9	40	54.1	187	10 O9LHK5	O9LHK5 arabidopsis
10	40	54.1	406	5 O19030	O19030 caenorhabdi
11	40	54.1	615	5 O18809	O18809 caenorhabdi
12	40	54.1	777	5 O95Y32	O95Y32 caenorhabdi
13	40	54.1	854	15 O12008	O12008 chimpanzee
14	40	54.1	863	15 O12007	O12007 chimpanzee
15	40	54.1	866	15 O12006	O12006 chimpanzee
16	40	54.1	868	15 O12009	O12009 chimpanzee

17	40	54.1	870	15 O12011	O12011 chimpanzee
18	40	54.1	1045	5 O9NKV1	O9NKV1 drosophila
19	39	52.7	41	12 O9LFC7	O9LFC7 chilo iride
20	39	52.7	101	2 O46552	O46552 bacteroides
21	39	52.7	228	10 O9FH99	O9FH99 arabidopsis
22	39	52.7	286	2 O9RLV6	O9RLV6 lactococcus
23	39	52.7	286	16 O9CIN5	O9CIN5 lactococcus
24	39	52.7	344	5 O18826	O18826 caenorhabdi
25	39	52.7	407	16 O9RK30	O9RK30 streptomyce
26	39	52.7	409	10 O9SR23	O9SR23 arabidopsis
27	39	52.7	505	10 O9CAY4	O9CAY4 arabidopsis
28	39	52.7	825	12 O9DTD4	O9DTD4 tt virus. o
29	39	52.7	1236	4 O9Y4C1	O9Y4C1 homo sapien
30	39	52.7	1854	11 O63131	O63131 rattus norv
31	39	52.7	2317	11 O63130	O63130 rattus norv
32	39	52.7	2338	11 O63132	O63132 rattus norv
33	38	51.4	70	5 O9BH73	O9BH73 conus texti
34	38	51.4	234	16 O99V31	O99V31 staphylococ
35	38	51.4	235	16 O9CDP1	O9CDP1 lactococcus
36	38	51.4	259	5 O46001	O46001 caenorhabdi
37	38	51.4	273	10 O94JN6	O94JN6 nicotiana t
38	38	51.4	275	2 O8VSN4	O8VSN4 shigella fl
39	38	51.4	346	5 O9U2L5	O9U2L5 caenorhabdi
40	38	51.4	405	2 O9AFX2	O9AFX2 shigella fl
41	38	51.4	406	10 O9LIL22	O9LIL22 zamia furfu
42	38	51.4	479	10 O9MAR8	O9MAR8 arabidopsis
43	38	51.4	482	12 O64837	O64837 human adeno
44	38	51.4	491	10 O9ZVA9	O9ZVA9 arabidopsis
45	38	51.4	493	10 O94JN7	O94JN7 nicotiana t

ALIGNMENTS

RESULT 1

P70860 ID P70860 PRELIMINARY: PRT; 425 AA.
AC P70860;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE FLGL protein.
GN FLGL.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID-139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-212;
RX MEDLINE-97312006; PubMed-9168617;
RA Ge Y., Old I.G., Girons I.S., Charon N.W.;
RT "The flgK motility operon of Borrelia burgdorferi is initiated by a
sigma 70-like promoter";
RL Microbiology 143:1681-1690(1997).
DR EMBL; U62901; AAB62741.1; -;
SQ SEQUENCE 425 AA; 47646 MW; 906AF88576FB703B CRC64;

Query Match 59.5%; Score 44; DB 2; Length 425;
Best Local Similarity 70.0%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 CRSWNKADNR 10
|||||
DB 355 CRSWSKKKNR 364

RESULT 2

O23598 ID O23598 PRELIMINARY: PRT; 760 AA.
AC O23598;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE DNA chromosome 4, ESSA I CONTIG fragment NO. 8.
GN A14G17570.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98121113; PubMed=9461215;
RA Bevan M., Bancroft I., Bent E., Love K., Goodman H., Dean C.,
RA Berkamp R., Dirkse W., van Staveren M., Stiekema W., Drost L.,
RA Ridley P., Hudson S.A., Patel K., Murphy G., Piffanelli P., Wedler H.,
RA Wedler E., Wambutt R., Weitzenecker T., Pohl T.M., Terryn N.,
RA Gielen J., Villarroel R., De Clerck R., van Montagu M., Lecharny A.,
RA Auborg S., Gy I., Kreis M., Lao N., Kavanagh T., Hempel S., Kottler P.,
RA Entian K.D., Rieger M., Schaeffer M., Funk B., Mueller-Auer S.,
RA Silvey M., James R., Montfort A., Pons A., Puigdomenech P., Douka A.,
RA Vouklatou E., Milioni D., Hatzopoulos P., Piravandi E., Obermaier B.,
RA Hilbert H., Duesterhoft A., Moores T., Jones J.D.G., Eneva T.,
RA Palme K., Benes V., Rechman S., Ansorge W., Cooke R., Berger C.,
RA Dalseny M., Voet M., Volckaert G., Mewes H.W., Klosterman S.,
RA Schueller C., Chalwatzis N.;
RT "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of
RT Arabidopsis thaliana";
RL Nature 391:485-488(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z97343; CAB10537.1; ..
DR EMBL; AL161546; CAB78760.1; ..
DR InterPro: IPR001810; F-box.
DR Pfam; PF00646; F-box; 1.
DR SMART; SM00256; FBOX; 1.
KW Hypothetical protein.
SQ SEQUENCE 760 AA; 86691 MW; B332B26B5C8F3A2B CRC64;

Query Match 58.8%; Score 43.5; DB 10; Length 760;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 8; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
OY 1 CRSWKADNRSC 12
DB 411 CRSWNSA-SRGC 421
|||||
RESULT 3
O9AMT8 PRELIMINARY; PRT; 445 AA.
AC O9AMT8;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE ID872 (EC 6.3.1.2).
GN ID872.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=110SPC4;
RX MEDLINE=21101824; PubMed=11157954;
RA Gottfert M., Rothlisberger S., Kundig C., Beck C., Marty R.,
RA Hennecke H.;
RT "potential symbiosis-specific genes uncovered by sequencing a 410-kb
RT DNA region of the Bradyrhizobium japonicum chromosome.";
RL J. Bacteriol. 183:1405-1412(2001).
CC -!- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) -> ADP + PHOSPHATE +
CC L-GLUTAMINE.
CC -!- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
DR EMBL; AF322013; AAG61059.1; ..

DR InterPro: IPR001637; GlnA_adenyltn.
DR InterPro: IPR001691; GlnA_synth.
DR Pfam: PF00120; gln-synt; 1.
DR ProDom; PD001057; GlnA_adenyltn; 1.
KW Ligase.
SQ SEQUENCE 445 AA; 48649 MW; AEA896E6A279388 CRC64;
Query Match 56.1%; Score 41.5; DB 2; Length 445;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 8; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
OY 1 CRSWKADNRSC 12
DB 314 CPSWG-VDNRSC 324
|||||
RESULT 4
O50464 PRELIMINARY; PRT; 262 AA.
ID O50464;
AC O50464;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Hypothetical protein RV1249c.
GN RV1249C OR MTV006.21C OR MT1287.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Horsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skellton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=CDC 1551 / OSHKOSH;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL021006; CAAL5905.1; ..
DR EMBL; AE007004; AAK45545.1; ..
DR TIGR; MT1287; ..
DR Tuberculist; RV1249c; ..
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 262 AA; 27571 MW; EC43F9CFA42C77BD CRC64;
Query Match 55.4%; Score 41; DB 16; Length 262;
Best Local Similarity 70.0%; Pred. No. 27;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 2 RSWNKADNR 11
DB 7 RSWKRFNRS 16
|||||
RESULT 5
O17729


```

ID O17729 PRELIMINARY; PRT; 477 AA.
AC O17729;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE D1086.6 protein.
GN D1086.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Smye R.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; 281491; CAB04023.2; -.
SQ SEQUENCE 477 AA; 53218 MW; 8263E004D61A231C CRC64;

Query Match 55.4%; Score 41; DB 5; Length 477;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 7; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

QY 1 CRSWNKADN--RSC 12
DB 96 CEKWNKCDVQRQC 109

RESULT 6
Q92GA5 PRELIMINARY; PRT; 1162 AA.
ID Q92GA5;
AC Q92GA5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical protein RC1218.
GN RC1218.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MALISH 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL Science 293:2093-2098(2001).
DR EMBL; AE008669; AAL03756.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1162 AA; 129191 MW; B60510E14D94B80D CRC64;

Query Match 55.4%; Score 41; DB 16; Length 1162;
Best Local Similarity 54.5%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CRSWNKADNRS 11
DB 65 CPPWNKSDGRN 75

RESULT 7
Q64736 PRELIMINARY; PRT; 2340 AA.
ID Q64736
AC Q64736;

```

```

DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE ROS1 PROTOONCOGENE (C-ROS PROTOONCOGENE).
GN ROS1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI; TISSUE=INTESTINE;
RX MEDLINE=95060820; PubMed=7970722;
RA Riethmacher D., Langholz O., Goedecke S., Sachs M., Birchmeier C.;
RT "Biochemical and functional characterization of the murine ros
RT protooncogene.";
RL Oncogene 9:3617-3626(1994).
RN [2]
RP SEQUENCE OF 1875-1973 FROM N.A.
RX MEDLINE=92380520; PubMed=1324876;
RA Narayana L., Nagarajan L.;
RT "A mouse c-ros genomic clone: identification of a highly conserved 22-
RT amino acid segment in the juxta-membrane domain.";
RL Gene 118:297-298(1992).
DR EMBL; M88610; AAA37463.1; -.
DR EMBL; M88609; AAA37463.1; JOINED.
DR EMBL; X81650; CAA57310.1; -.
DR HSSP; P11362; 1FGK.
DR MCD; MGI:97999; Rosl.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003962; FnIII-repeat.
DR InterPro; IPR003961; FN.III.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00041; fn3; 8.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00060; FN3; 5.
DR SMART; SM00135; LV; 2.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Repeat; Transferase; Tyrosine-protein kinase.
FT CONFLICT 13..13 K -> E (IN REF. 2).
FT CONFLICT 45..45 Q -> H (IN REF. 2).
SQ SEQUENCE 2340 AA; 261965 MW; A3A670B0C4151D7C CRC64;

Query Match 55.4%; Score 41; DB 11; Length 2340;
Best Local Similarity 41.7%; Pred. No. 2.6e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CRSWNKADNRSC 12
DB 62 CQFWNSVDQETC 73

RESULT 8
Q60705 PRELIMINARY; PRT; 2340 AA.
ID Q60705;
AC Q60705;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Proto-oncogene protein c-ros.
GN ROS1 OR C-ROS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-CD-1; TISSUE-KIDNEY;
RX MEDLINE-96130560; PubMed-8544427;
RA Kanwar Y.S., Liu Z.Z., Kumar A., Wada J., Carone F.A.;
RT "Cloning of mouse c-ros renal cDNA, its role in development and
RL relationship to extracellular matrix glycoproteins.";
DR Kidney Int. 48:1646-1659(1995).
DR EMBL; U15443; AAA50600.1; -.
DR HSSP; P11362; IFCF.
DR MGD; MGI:97999; Rosl.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003962; FN.II.repeat.
DR InterPro; IPR003961; FN.III.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00041; fn3; 9.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00014; FNTYPEII.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00060; FN3; 4.
DR SMART; SM00135; LY; 2.
DR SMART; SM00219; TYFKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Repeat; transferase; tyrosine-protein kinase.
SQ SEQUENCE 2340 AA; 262201 MW; 6FB82CC61C97E2EB CRC64;

Query Match 55.4%; Score 41; DB 11; Length 2340;
Best Local Similarity 41.7%; Pred. No. 2.6e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CRSWNKADNRSC 12
DB 62 CQFWSVDQETC 73

RESULT 9
Q9LHK5 PRELIMINARY; PRT; 187 AA.
AC Q9LHK5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Similarity to En/Spm-like transposon protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Nakamura Y.;
RT "Structural Analysis of Arabidopsis thaliana Chromosome 3. III.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP02043; BAB02612.1; -.
SQ SEQUENCE 187 AA; 21139 MW; B42A8FE4F9686E1 CRC64;

Query Match 54.1%; Score 40; DB 10; Length 187;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 WNKADNRSC 12
DB 122 WSKTDNLSC 130

RESULT 10
Q19030 PRELIMINARY; PRT; 406 AA.
AC Q19030;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

```

```

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 46.3 kDa protein.
GN E02C12.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE-99069613; PubMed-9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RL investigating biology. The C. elegans Sequencing Consortium.";
RT Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Pauley A., Gattung S.;
RT "The sequence of C. elegans cosmid E02C12.";
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U41995; AAA83463.1; -.
DR InterPro; IPR004119; DUF227.
DR Pfam; PF02958; DUF227; 1.
KW Hypothetical protein.
SQ SEQUENCE 406 AA; 46251 MW; D04CB4F4AABFCB45 CRC64;

Query Match 54.1%; Score 40; DB 5; Length 406;
Best Local Similarity 41.7%; Pred. No. 63;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CRSWNKADNRSC 12
DB 58 CRGWKEASSKVC 69

RESULT 11
Q18809 PRELIMINARY; PRT; 615 AA.
AC Q18809;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE C53D6.6 protein.
DE C53D6.6 protein.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gardner A.E.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99069613; PubMed-9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RL investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z70270; CAA94225.1; -.
SQ SEQUENCE 615 AA; 70011 MW; FF4C6F4C18EB5A75 CRC64;

Query Match 54.1%; Score 40; DB 5; Length 615;
Best Local Similarity 54.5%; Pred. No. 96;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```



```

O12006
ID O12006 PRELIMINARY; PRT; 866 AA.
AC O12006;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Envelope glycoprotein gp160.
GN ENV.
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ANT5;
RX MEDLINE=97184566; PubMed=9032368;
RA Nyambi P., Lewi P., Peeters M., Janssens W., Heyndrickx L.,
RA Fransen K., Andries K., Van den Haesevelde M., Heeney J., Piot P.,
RA Van Der Groen G.;
RT "Study of the dynamics of neutralization escape mutants in a
RT chimpanzee naturally infected with the simian immunodeficiency virus
RT SIVcpz-ant.";
RL J. Virol. 71:2320-2330(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ANT5;
RA Blouin J.C.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U84098; AAB58533.1; -.
DR InterPro: IPR000328; Env_GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 866 AA; 99464 MW; D7E49BDCB86372A6 CRC64;

Query Match 54.1%; Score 40; DB 15; Length 866;
Best Local Similarity 75.0%; Pred No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 WNKADNRS 11
Db 360 WNKVDNRT 367

```

Search completed: July 16, 2003, 13:15:16
 Job time : 34.4878 secs